

Fusion de données dans un système multicateurs pour l'étude du mouvement d'un objet dans une scène réelle

N. CHAFIQUI, A. BENSRAIR, P. MICHÉ

LCIA / INSA de Rouen
Place Emile Blondel B.P. 08
76131 Mont-Saint-Aignan cedex

Noufissa.Chafiqui@insa-rouen.fr

Résumé - Dans ce papier, nous présentons une nouvelle méthode de détermination des paramètres de rotation d'un objet sans connaissance a priori. Cette méthode exploite le fait que l'état d'une onde réfléchie est fonction de l'orientation de la surface observée. Une mise en correspondance spatio-temporelle fondée sur l'étude des histogrammes procure la valeur de l'angle de rotation autour d'un axe quelconque. Des résultats expérimentaux sur scènes réelles sont présentés.

Abstract - In this paper, we present a new method that enables the determination of rotation parameters of objects. This method is based on the fact that the polarisation state of a reflected beam is a function of the orientation of the observed surface. A spatio-temporal segmentation based on histogram analysis provides the rotation angle around any axis. Some experimental results about real scenes are presented.

1. Introduction

Le but principal de l'analyse de scènes dynamiques est de permettre l'extraction de l'information concernant directement l'environnement 3D dans lequel évolue le capteur. Les problèmes de l'analyse du mouvement 3D est plus particulièrement ceux de sa segmentation et de son estimation est un domaine très important en vision par ordinateur et en robotique mobile et concernent plusieurs domaines d'applications. Ces dernières incluent la navigation de robot, le suivi d'objets ou de cible, la surveillance dynamique, la segmentation de scènes complexes en vue de la compression d'images, l'imagerie satellitaire, le domaine biomédical ou la construction de modèles d'un environnement. Pour tous ces domaines, la qualité de l'estimation du mouvement est crucial et dépend en particulier des étapes de segmentation de mouvement ou la mise en correspondance spatio-temporelle [3].

Le principe utilisé dans cet article diffère des méthodes conventionnelles en ce sens qu'il exploite une propriété physique assez peu utilisée. Cette propriété est fondée sur l'analyse de l'état de polarisation des rayons lumineux réfléchis par la scène comme information sur l'orientation des surfaces des objets étudiés [4].

L'analyse de l'état de polarisation d'un faisceau lumineux réfléchi montre la possibilité de caractériser la forme d'un objet en trois dimensions de façon passive (sans aucune forme de contrôle de l'éclairage). Le procédé physique exploite le fait qu'un front d'ondes lumineuses prend un état de polarisation partiel particulier à la réflexion sur une surface. La principale caractéristique recherchée est l'angle existant entre une référence verticale et le plan de polarisation prédominant. Cependant, cette approche nécessite l'adjonction d'un élément optique sensible à l'état de polarisation des faisceaux incidents : un filtre polarisant dont l'orientation est commandée par un moteur pas à pas fournit une solution intéressante. Il apparaît donc que l'utilisation de l'analyse de

l'état de polarisation associée à un système de stéréovision procure des informations pertinentes pour l'analyse du mouvement en recherchant l'orientation des surfaces des objets à chaque nouvelle position.

2. Principe

Une onde lumineuse non polarisée présente, après réflexion sur une surface spéculaire, une polarisation rectiligne dont le degré dépend de l'angle d'incidence. On peut considérer l'onde réfléchie comme composée de deux ondes polarisées rectilignement respectivement parallèle et perpendiculaire au plan d'incidence. Les équations de Fresnel donnent les intensités de ces deux composantes $I_{//}$ et I_{\perp} , en fonction de l'angle d'incidence i et de l'angle de réfraction r :

$$\begin{cases} I_{//} = I_0 \frac{\operatorname{tg}^2(i-r)}{\operatorname{tg}^2(i+r)} \\ I_{\perp} = I_0 \frac{\sin^2(i-r)}{\sin^2(i+r)} \end{cases} \quad (1)$$

où I_0 désigne l'intensité de l'onde incidente.

L'étude de ces équations montre que l'intensité de la polarisation perpendiculaire au plan d'incidence I_{\perp} , en dehors des cas limites $i=0^\circ$ et $i=90^\circ$, est toujours supérieure à celle de la polarisation parallèle $I_{//}$. L'angle que forme le plan de polarisation prépondérant avec une référence verticale est appelé angle de polarisation. La détermination de cet angle permet de trouver l'orientation d'une surface.

3. Recherche de l'angle de polarisation

La détermination de l'orientation d'une surface ne peut donc être obtenue qu'après l'acquisition de l'angle de polarisation. Cependant, en traitement d'images, le capteur photosensible permet le passage d'une image codée sous forme photométrique à une image codée sous forme électrique. Un aspect essentiel de ces capteurs est qu'ils ne sont pas sensibles à l'état de polarisation de la lumière. Pour

résoudre ce problème, l'adjonction d'un processeur optique sensible à cette grandeur est nécessaire. La mesure de l'angle de polarisation passe alors par une modulation de l'intensité du faisceau lumineux en fonction de son état de polarisation.

L'onde issue de la réflexion sur une surface se compose de deux ondes dont les plans de polarisation sont respectivement parallèle et perpendiculaire au plan d'incidence. Quelle que soit la direction de l'incidence (sauf pour les cas limites) le plan perpendiculaire est toujours prédominant. Conformément à la loi de Malus, on peut écrire :

$$I_{\text{sortant}} = I_0 [\cos^2(\alpha - \theta) + \beta \cos^2(\alpha - \theta + \pi/2)] + \varepsilon \quad (2)$$

où I_0 représente l'intensité du champ perpendiculaire, α l'angle de polarisation recherché et qui représente l'angle entre ce plan prédominant et une référence verticale. β l'intensité relative du champ parallèle par rapport au champ perpendiculaire ($\beta \leq 1$). ε est un terme de bruit.

La méthode retenue pour la mesure des angles de polarisation pour chaque point d'une image est globale. Son principe consiste à effectuer des acquisitions d'images pour quatre positions angulaires du filtre polarisant. L'utilisation par la suite de la méthode des moindres carrés définit un algorithme qui permet de déterminer à partir de ces quatre intensités mesurées l'angle de polarisation recherché. Le critère des moindres carrés peut être écrit sous la forme :

$$J = \sum_{i=1}^4 (I_i - (I_0 \cos^2(\alpha - \theta_i) + \varepsilon'))^2 \quad (3)$$

où I_i désigne l'intensité de l'image i correspondant à une position θ_i du filtre polarisant et :

$$\begin{cases} I_0' = I_0(1 - \beta) \\ \varepsilon' = I_0\beta + \varepsilon \end{cases} \quad (4)$$

Un compromis entre la recherche de la précision de α et les contraintes technologiques et pratiques a conduit à travailler avec 4 images correspondant à des orientations $\theta = 0^\circ, 45^\circ, 90^\circ$ et 135° . Le choix de ces valeurs entraîne une simplification du calcul de α sous la forme suivante :

$$\alpha = \frac{\arctg\left(\frac{I_2 - I_4}{I_1 - I_3}\right)}{2} \quad (5)$$

4. Orientation d'une surface

Le modèle le plus adapté consiste à représenter une surface par des mosaïques de petits éléments plans [2]. La reconstruction de surface consiste à obtenir une adéquation entre ces plans et les grandeurs de polarisation. La caractérisation de ces plans passe par la définition d'un repère lié au système d'acquisition qui peut être étendu à tout l'espace pour définir un repère tridimensionnel.

En définissant le plan image du système d'acquisition par les axes x et y , et par l'axe z l'axe optique de la caméra, l'orientation d'un élément de surface dS , conforme à la

figure 1, est entièrement caractérisée par p et q , dérivées partielles de z par rapport aux deux axes x et y .

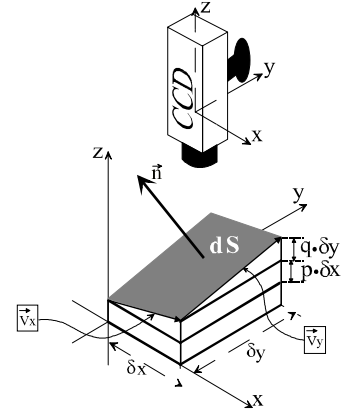


FIG. 1 : Représentation de l'orientation d'une surface par les pentes p et q

où p et q vérifient les relations suivantes :

$$\begin{cases} p = \frac{\partial z}{\partial x} \\ q = \frac{\partial z}{\partial y} \end{cases} \quad (6)$$

Les valeurs p et q permettent de déterminer d'autres grandeurs propres à l'élément de surface. Soit \vec{v}_x et \vec{v}_y les deux vecteurs générateurs du plan vectoriel associé à dS (cf. FIG 1). Après normalisation des composantes en x et en y , ces vecteurs peuvent s'écrire :

$$\begin{cases} \vec{v}_x = (1, 0, p)^T \\ \vec{v}_y = (1, 0, q)^T \end{cases} \quad (7)$$

Ainsi, le vecteur normal à l'élément de surface peut être obtenu à partir du produit vectoriel de \vec{v}_x et \vec{v}_y :

$$\vec{n} = \vec{v}_x \wedge \vec{v}_y = (-p, -q, 1)^T \quad (8)$$

Pour un élément de surface dS donné, la connaissance de ses pentes locales permet de déterminer le vecteur normal à cet élément. La détermination de l'ensemble de ces éléments plans et la recherche de leurs positions deviennent alors les principales étapes de l'estimation du mouvement d'un objet.

L'angle de polarisation que nous avons déterminé, permet, d'après la définition que nous en avons donné, de trouver le plan d'incidence en chaque point de l'image. Ce plan contient la normale à l'élément de surface objet associé au point image considéré. Cependant, cette connaissance est insuffisante pour obtenir la normale elle-même qui permet de trouver les pentes de l'objet à chaque instant. La solution de ce problème passe par la stéréovision. La scène est observée depuis deux points de vue distincts pour lesquels les orientations des plans d'incidence sont généralement différentes. Chacun de ces plans contient la normale à l'élément de surface recherchée. Celle-ci peut être obtenue en calculant l'intersection des deux plans d'incidence. L'application de ce principe pose une difficulté qui celle de connaître les points images stéréocorrespondants, provenant d'un même point objet. La solution de ce problème est obtenue par une fusion de données avec les résultats issus

d'un système de stéréovision photométrique classique [1].

La connaissance des angles de polarisation dans chacune des images pour un même point physique de la scène permet donc de définir les orientations des deux plans d'incidence se rattachant à ce point. Chacun de ces plans contient le vecteur normal à l'élément de surface centré sur le point considéré. Leur intersection est donc une droite vectorielle qui définit de façon unique le vecteur normal à l'élément de surface.

En utilisant les coordonnées de chaque vecteur, la normale à un élément de surface peut s'écrire :

$$\vec{n} = f \cdot \begin{pmatrix} -\sin(\alpha_g) \cdot x_{pd} \cdot \cos(\alpha_g) + \sin(\alpha_d) \cdot x_{pg} \cdot \cos(\alpha_g) \\ \sin(\alpha_d - \alpha_g) \cdot y_p - \cos(\alpha_d) \cdot \cos(\alpha_g) \cdot (x_{pg} - x_{pd}) \\ -f \cdot \sin(\alpha_d - \alpha_g) \end{pmatrix} \quad (9)$$

où x_{pd} et x_{pg} désignent les abscisses dans les plans image droit et gauche des projections d'un même point donné de la scène réelle, α_d et α_g représentent les angles de polarisation, augmenté de $\pi/2$, obtenus dans les images droite et gauche pour ce même point.

La configuration du système stéréo utilisée est particulière en ce sens que les éléments capteurs des deux caméras sont parfaitement coplanaires et alignés. Par conséquent, deux points stéréocorrespondants x_{pd} et x_{pg} possèdent une coordonnée commune y_p .

Il a été montré précédemment par l'équation (8) qu'il est possible de transposer la connaissance du vecteur normal à celle des pentes tridimensionnelles élémentaires. Cette équation amène les expressions de p et q en fonction des paramètres de polarisation et de positions définis par l'équation (10) :

$$\begin{cases} p = \frac{-\sin(\alpha_g) \cdot x_{pd} \cdot \cos(\alpha_g) + \sin(\alpha_d) \cdot x_{pg} \cdot \cos(\alpha_g)}{f \cdot \sin(\alpha_d - \alpha_g)} \\ q = \frac{\sin(\alpha_d - \alpha_g) \cdot y_p - \cos(\alpha_d) \cdot \cos(\alpha_g) \cdot (x_{pg} - x_{pd})}{f \cdot \sin(\alpha_d - \alpha_g)} \end{cases} \quad (10)$$

Nous avons donc deux cartes supplémentaires : la carte des pentes p et la carte des pentes q. En utilisant conjointement ces deux dernières données, nous pouvons décrire avec robustesse les différentes zones de la scène. Dans le cas d'objets polyédriques, ces zones correspondent aux facettes le constituant. L'information que nous obtenons sur la scène est globale, ce résultat nous a donc amené à choisir des primitives de types facettes planes. Ces facettes sont obtenues par la recherche de zones de recouvrement des régions extraites dans les deux cartes. Pour obtenir ces régions, une méthode de segmentation en région est appliquée sur les N images de la séquence. L'utilisation des facettes permet d'intégrer très naturellement les concepts de surface et des normales.

5. Extraction de facettes planes

L'extraction des facettes nécessite des traitements intermédiaires. Rappelons ci-dessous les principales étapes de ce processus :

❶ L'image est partitionnée en régions homogènes représentant les pentes p et q. Ces pentes sont supposées être

la projection d'une même facette dans l'espace 3D.

❷ Les primitives extraites dans les images des pentes p et q pour un même objet sont mises en correspondance deux à deux pour former des facettes.

❸ La dernière étape consiste à sélectionner les facettes planes possédant une cohérence physique. Autrement dit, éliminer celles dues au bruit, à une erreur de segmentation ou encore à une erreur de calcul des pentes. Nous utiliserons lors de cette étape une méthode ensembliste en comparant l'image à un élément structurant. Une transformation par dilatation érosion établie avec cet élément structurant nous permet d'extraire les principales facettes planes de l'objet.

Ainsi, nous avons représenté chaque objet présent dans la scène par ses facettes planes ; pour chaque facette plane, nous disposons des coordonnées de son vecteur normal. Ce résultat constitue une information pertinente pour l'étude de la rotation.

En effet, dans le cas de la rotation d'un objet dans une scène, nous définissons par $\vec{n}_{i,1}$ le vecteur normal à la facette i à l'instant 1. La détermination des nouvelles coordonnées du vecteur normal à l'instant 2 permet de situer dans l'espace le vecteur $\vec{n}_{i,2}$.

6. Rotation autour d'un axe quelconque

Le problème qui se pose est l'appariement des vecteurs normaux à la facette i aux instants 1 et 2, il s'agit donc d'établir une mise en correspondance spatio-temporelle des coordonnées des vecteurs normaux.

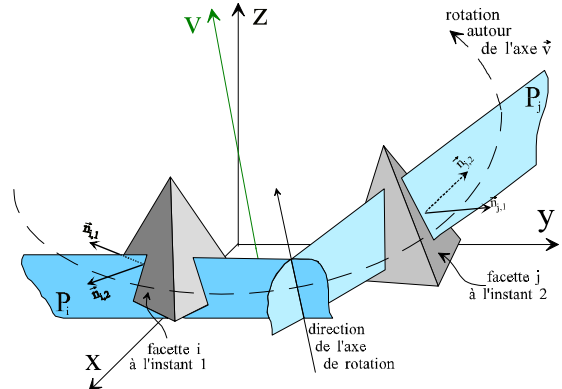


Figure 3 : Plans générés par deux facettes en rotation

La mise en correspondance spatio-temporelle d'au moins deux facettes fait logiquement apparaître deux plans distincts. Chacun de ces deux plans contient l'axe de rotation \vec{v} , l'intersection de ces deux plans est donc une droite vectorielle qui donne l'axe de rotation sous la forme :

$$\vec{v} = [(\vec{n}_{i,1} + \vec{n}_{i,2}) \wedge (\vec{n}_{i,1} \wedge \vec{n}_{i,2})] \wedge [(\vec{n}_{j,1} + \vec{n}_{j,2}) \wedge (\vec{n}_{j,1} \wedge \vec{n}_{j,2})] \quad (11)$$

Le calcul de l'axe de rotation permet ensuite de donner l'angle de rotation α sous la forme :

$$\alpha = \arccos \left(\frac{\vec{u}_{i,1} \cdot \vec{u}_{i,2}}{\|\vec{u}_{i,1}\| \|\vec{u}_{i,2}\|} \right) \quad (12)$$

avec :

$$\begin{cases} \vec{u}_{i,1} = \vec{n}_{i,1} \wedge \vec{v} \\ \vec{u}_{i,2} = \vec{n}_{i,2} \wedge \vec{v} \end{cases} \quad (13)$$

Ainsi, pour chaque combinaison de couple de

facettes $\{(F_{i,1}, F_{i,2})\} (F_{j,1}, F_{j,2})$ un angle et un axe de rotation sont proposés

7. Résultats

Nous avons appliqué notre méthode à la détermination de la rotation d'une pyramide de taille approximative de 3 cm. L'objet est placé sur un plan horizontal situé à 15 cm du système d'acquisition. Entre les deux instants les conditions d'acquisition sont sensiblement identiques. Le système d'acquisition est constitué de deux caméras coplanaires. Les images sont au format 512x512 pixels. Le temps de traitement nécessaire à l'élaboration de la carte des facettes est de l'ordre de 6s (sur un Pentium 166Mhz) sans tenir compte des phases d'acquisition des cartes d'angles de polarisation. Les résultats seront présentés en trois parties :

- ❶ les deux premières images concernent le résultat de calcul des pentes,
- ❷ la troisième image représente la carte des facettes de l'objet, cette image sera suivie d'une présentation des coordonnées de chaque facette.

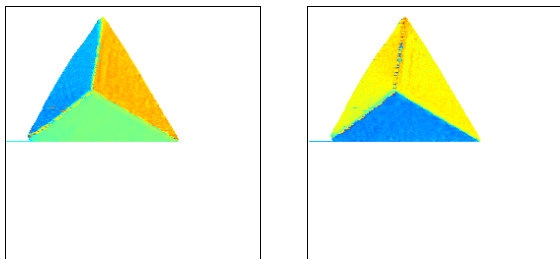


FIG 4 : Cartes des pentes p et q à l'instant 1

Une mise en correspondance spatiale des cartes segmentées procure l'image des facettes suivante :

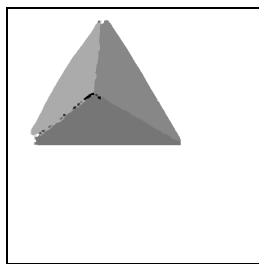


FIG 5 : Facettes extraites à l'instant 1

Dans la figure ci-dessus, trois facettes planes notées $F_{1,1}$, $F_{2,1}$ et $F_{3,1}$ sont extraites :

$$F_{1,1} \Rightarrow \begin{cases} M_p^*(1) = -68 \\ M_q^*(1) = 41 \end{cases} \quad F_{2,1} \Rightarrow \begin{cases} M_p^*(2) = -1 \\ M_q^*(2) = -78 \end{cases} \quad F_{3,1} \Rightarrow \begin{cases} M_p^*(3) = 63 \\ M_q^*(3) = 41 \end{cases}$$

Une rotation de 11° de la pyramide est effectuée autour de l'axe z. Les figures suivantes donnent le résultat de cette opération :

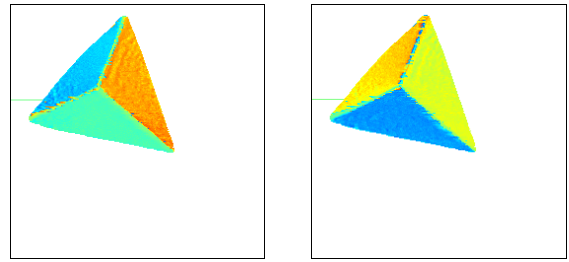


FIG 6 : Cartes des pentes p et q à l'instant 2

Trois nouvelles facettes sont extraites :

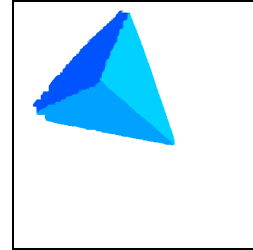


FIG 7 : Facettes extraites à l'instant 2

$$F_{2,2} \Rightarrow \begin{cases} M_p^*(2) = -16 \\ M_q^*(2) = -77 \end{cases} \quad F_{3,2} \Rightarrow \begin{cases} M_p^*(3) = 69 \\ M_q^*(3) = 28 \end{cases} \quad F_{3,1} \Rightarrow \begin{cases} M_p^*(3) = 63 \\ M_q^*(3) = 41 \end{cases}$$

Une fois résolu le problème de la mise en correspondance temporelle, un angle et un axe de rotation sont proposés à partir des équations 11, 12 et 13, pour chaque combinaison de couple de facettes $\{(F_{i,1}, F_{i,2})\} (F_{j,1}, F_{j,2})$.

Les trois combinaisons possibles de couples de facettes procurent un angle de rotation $\alpha \in [11^\circ, 11.97^\circ]$ autour d'un axe dont la direction correspond à z.

8. Conclusion

Les résultats présentés ci-dessus montrent la capacité à estimer l'axe et l'angle de rotation avec une précision dépassant les 90%, même si quelques problèmes liés à une trop grande ressemblance polarimétrique entre les deux points de vue de la scène engendre une légère erreur de la détermination des pentes locales de l'objet. L'approche vectorielle adoptée ne permet pas pour l'instant de situer le centre de rotation. Pour déterminer ce centre, une solution consiste à rechercher, pendant le processus d'extraction de facettes planes, son centre d'inertie C_i . La détermination de ce centre à l'instant suivant permettra une recherche ce centre instantané de rotation CIR. Cette méthode, basée sur un suivi de points, sera sensible aux erreurs de calcul. Le centre de rotation sera donné dans l'espace 3D comme appartenant à une sphère dont le rayon devra être déterminé.

9. Bibliographie

- [1] A. BENSRAIR, & al : "Fast and automatic stereo vision matching algorithm based on dynamic programming method". *Pattern Recognition Letters*, 17, 1996, pp. 457-466.
- [2] B.K.P HORN "Image intensity understanding" *Artificial Intelligence*, vol. 8, n°2, p. 201-231, 1977.
- [3] H. L. LI, "Motion estimation of two-dimensional objects based on the straight line hough transform : a new approach" *Pattern recognition* vol. 29, n° 8, 1996, p. 1245-1258.
- [4] L. B. WOLFF & al, "Constraining object features using a polarization reflectance model", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.13 (7), pp.635-657, 1991.