

DETERMINATION DE LA COMPOSANTE AXIALE DU MOUVEMENT A PARTIR D'UNE SEQUENCE D'IMAGES STEREO

*Stéphane MOUSSET, Pierre MICHÉ,
Roland DEBRIE*

*La3I/LCIA INSA de Rouen BP 08 76131 Mont Saint Aignan Cédex
Tél : 35 52 84 07 Fax : 35 52 84 83*

RÉSUMÉ

Dans la présente communication, nous présentons une méthode robuste et rapide permettant le calcul en temps réel de la composante axiale de la vitesse des différents points d'une scène à partir d'une séquence d'images stéréoscopiques.

L'application de cette méthode pose trois problèmes auxquels nous apportons des solutions. Le premier problème est lié à la réduction d'un bruit vertical spécifique, le second au calcul de la vitesse des objets éloignés, et le dernier à la mise en correspondance spatio-temporelle.

Des résultats expérimentaux obtenus sur des images réelles sont présentés à la fin de ce papier.

ABSTRACT

In this paper, we present a robust and fast method that enables real-time computing of axial motion from stereo image sequences.

The application of this method causes three problems that we have solved. The first problem is in relation with reducing a specific vertical noise due the stereovision process, the second is about computing motion of far objects, and the last is about spatio-temporal matching.

Experimental results obtained on several real stereo images are given at the end of this paper.

1. Introduction

La détermination du mouvement d'objets à partir de scènes réelles est une étape essentielle dans le fonctionnement des systèmes de guidage et de sécurité des véhicules autonomes.

La plupart des méthodes de détermination de la vitesse 3-D des objets réels procèdent en deux temps.

Tout d'abord il est nécessaire de réaliser une modélisation en trois dimensions des objets de la scène. Cela peut être obtenu soit à l'aide d'une seule caméra avec un déplacement relatif connu [1] [2], soit avec plusieurs caméras dont on connaît les positions relatives et dont les images sont mises en correspondance [3] [4] [5].

Cette reconstruction 3-D est suivie d'une mise en correspondance spatio-temporelle de primitives liées aux objets sur lesquelles porte le calcul de la vitesse. Le choix de ces primitives est un point important qui influe sur le volume des calculs à effectuer et sur la précision des résultats.

L'approche de Zang et Faugeras [4] privilégie des primitives de niveau intermédiaire, telles que des segments de droite, ce qui permet d'exploiter des contraintes objets dans la modélisation de l'espace 3-D.

Sabata and Aggarwal [6] font remarquer que le choix de points en tant que primitives pose des problèmes de bruit car ces éléments caractéristiques sont de bas niveau.

Les résultats obtenus par ces méthodes sont souvent intéressants, mais la reconstruction 3-D et la mise en correspondance spatio-temporelle requièrent d'importants moyens de calcul guère compatibles avec un système embarqué.

2. Détermination de la composante axiale de la vitesse

La méthode que nous avons développée est destinée à des applications d'assistance à la conduite de véhicules automobiles. Dans la présente communication nous ne présentons que la détermination de la composante axiale de la vitesse - suivant l'axe optique du système de prise de vue - qui est un des éléments les plus déterminants des applications visées.

2.1. Calcul de la vitesse

Considérons un point dont on peut déterminer les distances z_1 et z_2 par rapport au plan de prise de vue aux



instants t_1 et t_2 ; la composante axiale de sa vitesse s'exprime de manière immédiate par la relation suivante :

$$\text{Vitesse} = \frac{z_2 - z_1}{t_2 - t_1} \quad (1)$$

Le principe de notre méthode repose sur l'application cette relation en déterminant les distances par stéréovision passive. Pratiquement, nous avons utilisé un système de vision 3-D développé dans notre laboratoire capable d'établir des cartes de disparité en 0,1 seconde [7] [8]. Rappelons que la disparité de deux points image stéréocorrespondants est la différence de leurs coordonnées dans les deux images.

Tenant compte de ce que la distance d'un point objet au plan de prise de vue est une fonction inverse de la disparité de ses points image, la relation (1) s'exprime de la façon suivante :

$$\text{Vitesse} = \frac{K}{t_2 - t_1} \left(\frac{1}{\text{disp}(t_2)} - \frac{1}{\text{disp}(t_1)} \right) \quad (2)$$

où K est un paramètre dépendant des caractéristiques du système d'acquisition des images stéréoscopiques et $\text{disp}(t)$ la disparité d'un point objet à l'instant t.

La mise en oeuvre de cette relation utilise les données de plusieurs images. L'ensemble des opérations pour la détermination de la composante axiale de la vitesse est résumé par la figure 1 présentée ci-dessous.

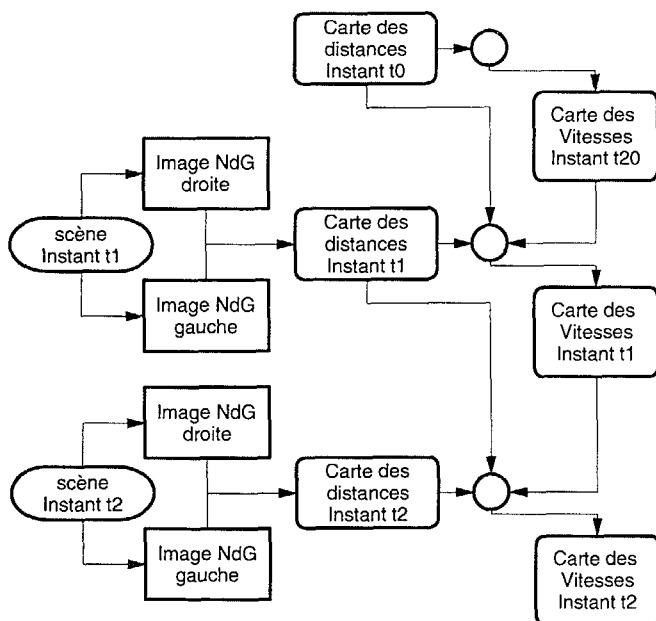


Figure 1 : Organisation des calculs pour l'obtention de la vitesse à l'instant t_2 .

2.2. Modélisation de l'environnement

Partant des cartes de disparité, nous avons simplifié la modélisation de l'espace 3-D, en considérant, non pas

des volumes, mais des plans d'iso-disparité perpendiculaires à l'axe optique du système de prise de vue. Ceci permet de s'affranchir des contraintes de non-déformation des solides couramment utilisées dans la reconstruction 3-D et la mise en correspondance spatio-temporelle. Par contre se pose le problème de la continuité. Nous exposons plus loin comment cette difficulté a été traitée.

3. Calcul de la carte des vitesses

Entre deux cartes de profondeur déterminées en temps réel à un rythme de 5 images par seconde, les déplacements des objets sont faibles. Il est alors possible de simplifier le calcul du mouvement par linéarisation des équations.

Connaissant les positions des plans aux instants t_1 et t_2 , il suffit théoriquement de dériver ces positions en fonction du temps pour obtenir la vitesse axiale. L'application de cette opération pose trois problèmes essentiels, auxquels nous apportons les solutions présentées ci-dessous.

3.1. Réduction de l'effet du bruit

Pour obtenir les cartes de profondeur nous utilisons un système stéréoscopique ayant une géométrie particulière et des algorithmes rapides adaptés à cette dernière [8]. Ces conditions particulières de calcul créent un bruit vertical spécifique. Une carte de vitesses calculée sans traitement du bruit donne une image inexploitable.

L'utilisation d'un filtre passe-bas usuel ou d'un filtre médian améliore les cartes de disparité, mais insuffisamment pour pouvoir les exploiter pour les calculs de la vitesse.

La réduction du bruit est réalisée par un traitement intégré dans le processus de calcul des cartes de vitesses. Ce traitement se décompose en trois étapes que nous exposons ci-dessous : prédiction, comparaison et validation.

La distance d'un plan objet à l'instant $t + \Delta t$ est prédite en fonction de sa vitesse et de sa distance à l'instant t. Pour cela, on utilise l'hypothèse que durant l'intervalle de temps Δt le mouvement est rectiligne et uniforme. Celle-ci est d'autant mieux vérifiée que l'intervalle Δt est faible.

La différence entre la disparité prédite et la disparité mesurée à l'instant $t + \Delta t$ pour un point image doit être inférieure à un seuil de tolérance variable en fonction de la disparité.

Si la distance est importante, ce qui se traduit par une disparité faible, la source d'erreur est due à une imprécision sur la disparité. Le seuil de tolérance est



calculé en estimant que la précision en disparité est de plus ou moins une unité.

Par contre, dans le cas de distances faibles, ce qui correspond à des disparités importantes, le seuil de tolérance est calculé par la considération d'une contrainte physique selon laquelle l'accélération d'un plan est limitée.

Dans la dernière étape, la vitesse des points image vérifiant les conditions énoncées ci-dessus est calculée. Les autres points sont marqués pour un traitement ultérieur. Ce traitement par prédiction a l'avantage de donner des valeurs précises de vitesse, mais il n'est pas adapté à certains cas particuliers que nous abordons ci-dessous.

3.2. Traitement des objets lointains

Dans le cas d'objets situés à des distances importantes du plan de prise de vue, un déplacement axial ne se traduit pas forcément par une variation de la disparité entre deux acquisitions consécutives. Pour résoudre ce problème, l'intervalle de temps utilisé dans le calcul de la vitesse n'est pas constant, mais est conditionné par une modification de la disparité sur les cartes de profondeur. Il en résulte que, pour une image donnée, les plans objet peuvent être traités avec des intervalles de temps différents.

Après ce traitement on retrouve les valeurs de l'incertitude obtenue en différenciant la relation (2) :

$$\frac{\Delta \text{Vitesse}}{\text{Vitesse}} = \frac{(\text{disp}(t) + \text{disp}(t + n \cdot \Delta t)) \cdot \Delta \text{disp}}{\text{disp}(t) \cdot \text{disp}(t + n \cdot \Delta t)} \quad (4).$$

où n est le paramètre d'échantillonnage temporel variable et Δdisp est fixé à 1. L'incertitude sur la vitesse ainsi obtenue diminue de façon importante avec l'augmentation de la disparité.

3.3. Recouvrement et non-recouvrement

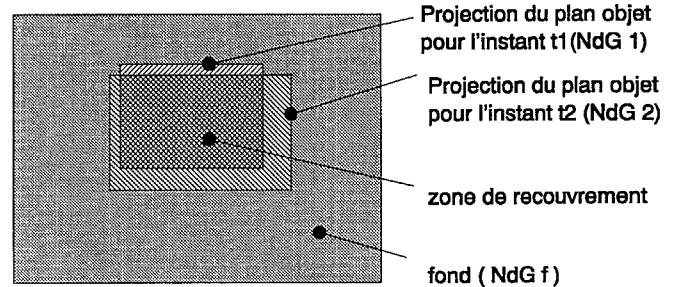
L'utilisation d'un algorithme de mise en correspondance spatio-temporel complique les calculs et ne nous permet plus d'obtenir les vitesses axiales en temps réel. Il est nécessaire d'utiliser une méthode pour contourner ce problème.

La fréquence d'acquisition des images est élevée. Cela permet de faire l'hypothèse que les plans objets se déplacent peu entre deux acquisitions. Il y a alors recouvrement des surfaces projetées entre les deux cartes de profondeur. C'est dans cette zone que l'on calcule la vitesse. Le résultat doit être alors étendu à l'ensemble de la surface projetée de l'instant t2.

Nous avons décidé d'utiliser une transformation morphologique que l'on appelle dilatation géodésique en niveaux de gris et que l'on peut résumer par la figure (2).

Cette transformation utilise comme primitive le point, le temps de calcul dépend donc de la taille de l'image et non du nombre d'objets à traiter. Elle permet aussi de faire respecter la condition de continuité des objets.

superposition des cartes de profondeur entre les instants t1 et t2



Transformation pour obtenir la carte des vitesses à l'instant t2

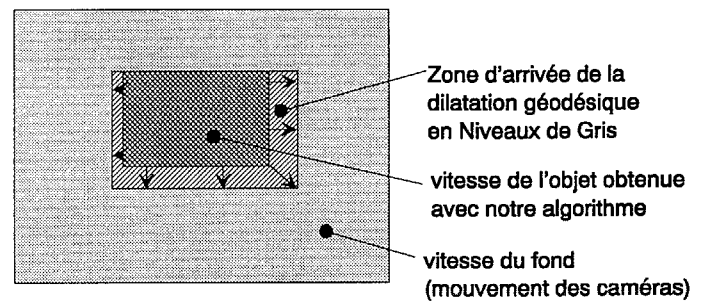


Figure 2 : Dilatation géodésique en niveaux de gris

La zone d'arrivée de la dilatation est la projection de l'objet au nouvel instant. Les problèmes de disparition ou d'apparition d'objets dans le champ du système de prise de vue ne se posent pas.

4. Résultats et conclusion

Notre algorithme a été testé sur plusieurs images réelles 512 x 512 x 8 bits. Les résultats obtenus sont encourageants malgré des cartes de profondeurs bruitées, ce qui confirme la robustesse de notre algorithme.

Le tableau ci-dessous illustre les résultats. Après une période transitoire, la vitesse est donnée en temps réel avec une précision d'environ 10%.

Instant		t1	t2	t3	t4	t5	t6
Disparité moy.		62	66	72	78	86	94
Distance moy.	m	8,7	8,1	7,5	6,9	6,25	5,7
Vitesse réelle	m/s	?	5,2	6	5,9	6,2	5,7
Vitesse calculée	m/s	?	?	3,3	4,5	5,4	5,4

Tableau 1 : Résultat d'un essai

Dans un futur proche, le bruit vertical, lié à la non-cohérence interligne doit être diminué avec l'amélioration des algorithmes d'obtention de cartes de profondeur.



Signalons enfin qu'un système spécifique comportant 16 DSP est en cours de développement, grâce auquel le temps d'établissement d'une carte de vitesses, obtenu par simulation, sera de l'ordre de 0,2 s.

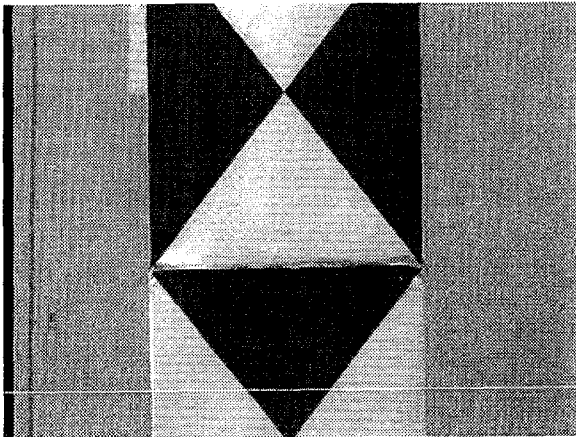
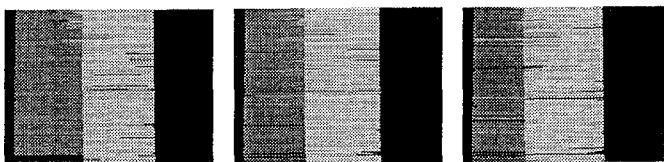


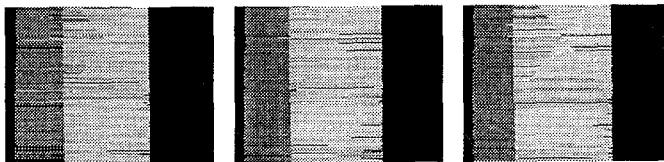
image de la scène réelle.



instant t1

instant t2

instant t3

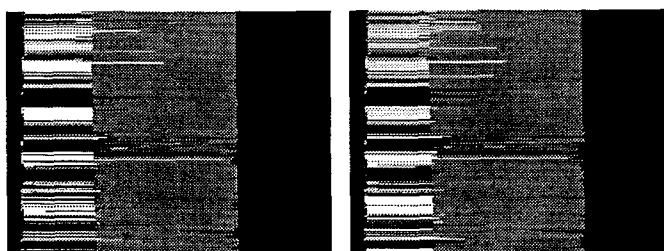


instant t4

instant t5

instant t6

Cartes de profondeur



instant t5

instant t6

cartes des vitesses obtenues par notre méthode.

Références bibliographiques

[1] D. KALIVAS and A. SAWCHUCK "A Region Matching Motion Estimation Algorithm", *CVGIP : IMAGE UNDERSTANDING*, 1991, Vol 54, N° 2, September, pp. 275-288.

[2] Y. LIU and T. HUANG "Three-Dimensional Motion Determination from Real Scene Images using Straight Line Correspondences", *Pattern Recognition*, 1992, Vol 25, N° 6, pp. 617-639.

[3] H. CHEN "Determining Motion and Depth from Binocular Orthographic Views", *CVGIP : IMAGE UNDERSTANDING*, 1991, Vol 54, N° 1, July, pp. 47-55.

[4] Z. ZHANG and O. FAUGERAS "Determining Motion from 3D Line Segment Matches : a Comparative Study", *IMAGE AND VISION COMPUTING*, 1991, Vol 9, N° 1, February.

[5] S. FOGEL "The Estimation of Velocity Vector Fields from Time-Varying Image Sequences.", *CVGIP : IMAGE UNDERSTANDING*, 1991, Vol 53, N° 3, May, pp. 253-287.

[6] B. SABATA and J.K. AGGARWAL "Estimation of Motion from a Pair of Range Images : A Review", *CVGIP : IMAGE UNDERSTANDING*, 1991, Vol 54, N° 3, November, pp 309-324.

[7] T. QUIGUER, P.MICHÉ, R. DEBRIE "Méthode d'extraction de primitives par seuillage automatique", 13ème colloque GRETSI sur le traitement du signal et des images, 16-20 septembre 1991, JUAN-LES-PINS.

[8] A. BENSHAIR, P.MICHÉ, R. DEBRIE "New Fast Stereo Matching for Parallel Processing", 2nd IEEE International Workshop on Advanced Motion Control, march 16-18 1992, NAGOYA, JAPAN