

UN NOUVEL ALGORITHME DE MISE EN CORRESPONDANCE
STEREO UTILISANT LA PROGRAMMATION DYNAMIQUE ET
DES PARAMETRES AUTO-ADAPTATIFS

*Abdelaziz BENSRAHAI, 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

ABSTRACT

Pour la réalisation d'un capteur de vision 3-D temps réel, nous avons développé un nouveau processus de stéréovision passive implémentable sur une architecture spécialisée. Ce système est destiné au guidage de robots mobiles ou à la conduite automatique de véhicules.

Ces applications nécessitent des algorithmes rapides, utilisant des paramètres décisionnels auto-adaptatifs.

Dans cette communication, nous proposons un nouvel algorithme de mise en correspondance, vérifiant ces critères, fondé sur une technique de programmation dynamique.

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

To realise a 3-D vision sensor, we have developed a new passive stereovision system suitable for implementation on specialised processors. This system can be applied to navigable robot and autonomous vehicles.

These applications require fast and self-adaptive algorithms.

In this paper, we present a new fast stereo matching algorithm based on a new dynamic programming method using self-adaptive decision parameters.

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

1. Introduction

Le principe de la stéréovision passive consiste à calculer par triangulation la position tridimensionnelle d'un point à partir des positions de ses projections dans deux images prises simultanément par deux caméras, dont les positions spatiales ont été déterminées lors d'une phase de calibration.

Tout d'abord des éléments caractéristiques sont extraits de chaque image, puis sont mis en correspondance, ce qui revient à identifier entre les deux images des informations bidimensionnelles homologues provenant d'un même élément physique de l'espace 3-D.

Pour résoudre le problème de la mise en correspondance, la programmation dynamique a été utilisée dans un premier temps par Baker et Benard. Ils appartiennent les points de contour entre une paire de lignes épipolaires conjuguées. Cette approche repose sur la recherche d'un chemin optimal dans un graphe bidimensionnel en minimisant une fonction coût, et en utilisant la contrainte d'ordre des primitives [1] [2].

Ohta et Lloyd ont repris cette méthode mais en recherchant un chemin optimal non seulement en intra-ligne mais aussi en inter-ligne, c'est-à-dire une recherche dans un graphe 3-D [3] [4].

En 89, Wu a développé un algorithme de mise en correspondance combinant la programmation dynamique et la corrélation [5].

Toutes les approches exposées ci-dessus concernent la reconstruction 3-D des objets de la scène, c'est pour cette raison qu'elles sont souvent très "gourmandes" en temps machine : plusieurs minutes de temps CPU sur un ordinateur Micro Vax 750. Par conséquent elles sont inopérantes quand il s'agit d'aider un robot dans son déplacement.

Toutefois, la PD présente un grand avantage par son aspect global de la recherche des correspondances entre primitives, nous pensons donc que cette méthode peut être retenue pour résoudre le problème de la mise en correspondance d'une façon optimale.

2. Notre approche

Le but de nos travaux est la réalisation d'un capteur de vision 3-D temps réel. Ceci nécessite des algorithmes simples, rapides et robustes, utilisant des paramètres de décision auto-adaptatifs. Pour répondre à ces critères, nous avons retenu une configuration de caméras (décrite ci-après) permettant de réduire les traitements de segmentation et de mise en correspondance à une seule



dimension. De plus, nous nous sommes attachés à développer un processus de stéréovision facilement parallélisable et qui pourra, à terme, être implémenté dans une architecture spécialisée.

2.1. Configuration des deux caméras

Pour simplifier les processus de segmentation et de la mise en correspondance, nous avons placé les deux caméras de prise de vues de telle sorte que leur capteur soient parfaitement alignés dans un même plan [6].

Dans cette configuration, les lignes épipolaires sont confondues avec les lignes de balayage horizontales des caméras et, de ce fait, les éléments correspondants des images droite et gauche se trouvent sur des lignes horizontales de même indice. Sachant cela, l'algorithme de mise en correspondance, peut être accéléré en réduisant le traitement à la seule dimension horizontale. Les attributs utilisés pour cette mise en correspondance peuvent donc ne dépendre que de cette seule dimension, et conséquemment, la segmentation elle-même peut être établie ligne par ligne.

2.2. Segmentation des deux images stéréos

La phase de segmentation de notre système de stéréovision utilise un algorithme rapide et auto-adaptatif d'extraction de points de contours pertinents. Ce traitement s'opère ligne par ligne pour chaque image stéréo et il est fondé sur un nouveau concept que nous avons baptisé déclivité [7]. Ce concept est défini par la variation de niveau de gris entre deux pixels correspondant à des extrema locaux adjacents. L'extraction des éléments caractéristiques est obtenue par un simple seuillage. La valeur du seuil est calculée pour chaque ligne image par une analyse statistique des déclivités, procurant l'auto-adaptivité au processus.

2.3. Définition des attributs de chaque primitive

Pour la recherche de correspondance, une structure de données est constituée à partir des éléments caractéristiques issus de la segmentation, à chacun desquels sont associés les attributs suivants :

- son abscisse sur la ligne image (qui servira pour le calcul de la disparité),
- les niveaux de gris des trois voisins gauches ($ndg[k]$ et $ngg[k]$ respectivement pour des éléments caractéristiques droit et gauche) ce qui caractérise le voisinage gauche de l'élément considéré,
- les niveaux de gris des trois voisins droits ($ndd[k]$ et $ngd[k]$ respectivement pour des éléments caractéristiques droit et gauche) représentant le voisinage droit de l'élément considéré. k variant de 1 à 3.

2.4. Définition des distances photométriques :

A partir des attributs photométriques ci-dessus, nous définissons les distances photométriques d'appariements de la manière suivante :

$$\text{distphg} = \sum_{k=1}^3 |ndg(k) - ngg(k)|$$

et

$$\text{distphd} = \sum_{k=1}^3 |ndd(k) - ngd(k)|$$

où distphg et distphd sont respectivement les distances photométriques gauche et droite, celles-ci permettent de mesurer la vraisemblance photométrique entre deux éléments candidats à l'appariement.

3. Algorithme de mise en correspondance

Pour la mise en correspondance, les méthodes classiques - que nous avons citées en introduction - fondent leurs algorithmes sur un critère de minimisation d'une fonction coût pour trouver un chemin optimal dans un graphe. En ce qui nous concerne, nous avons mis au point un algorithme qui utilise une nouvelle méthode de programmation dynamique basée sur un critère de maximisation d'une fonction gain.

Nous considérons un graphe bi-dimensionnel où l'axe horizontal représente la droite épipolaire gauche et l'axe vertical représente l'épipolaire droite correspondante. Chaque noeud du graphe représente un appariement possible entre un point pertinent droit et un point pertinent gauche auquel nous attribuons un gain local. A partir d'une valeur seuil g_{\max} , ce gain décroît en fonction des distances photométriques définies ci-dessus de la manière indiquée dans la figure 1. Dans le cas où le gain local s'annule, l'hypothèse de mise en correspondance considérée est rejetée. Ceci présente un avantage par rapport à la fonction coût qui peut augmenter indéfiniment et par conséquent ne rejette aucune possibilité d'appariement dans la première phase du traitement.

Ensuite, un certain nombre de chemins dans le graphe sont générés, en respectant la contrainte d'ordre des primitives, en attribuant à chaque chemin du graphe un gain cumulé G . Ce gain est défini lorsqu'on passe d'un état $(i-1, u(i-1))$ à l'état (i, j) de la façon suivante :

$$G(i,j) = \text{Max} [G(i-1,u(i-1)) + g(i-1,u(i-1),i,j)]$$

$$1 \leq u(i-1) \leq nc$$

où $g(i-1,u(i-1),i,j)$ est le gain élémentaire, autrement dit le gain d'un pas, associé à un arc dans le graphe, et nc est le nombre de points caractéristiques candidats à l'appariement pour un point caractéristique donné.

Les meilleurs appariements ou noeuds optimaux sont ceux qui constituent le chemin correspondant au gain cumulé maximum [8], un exemple illustrant cet aspect est donné dans la figure 2. Dans cette figure le chemin optimal est représenté par les cases doublés.

Quant au paramètre g_{max} que nous utilisons, il est calculé en fonction des seuils auto-adaptatifs issus de la phase de segmentation. Il est donc calculé pour chaque ligne épipolaire et nous permet de limiter considérablement le nombre d'hypothèses d'appariement d'une façon automatique. Ceci a pour conséquence d'atteindre une optimisation intéressante de l'algorithme.

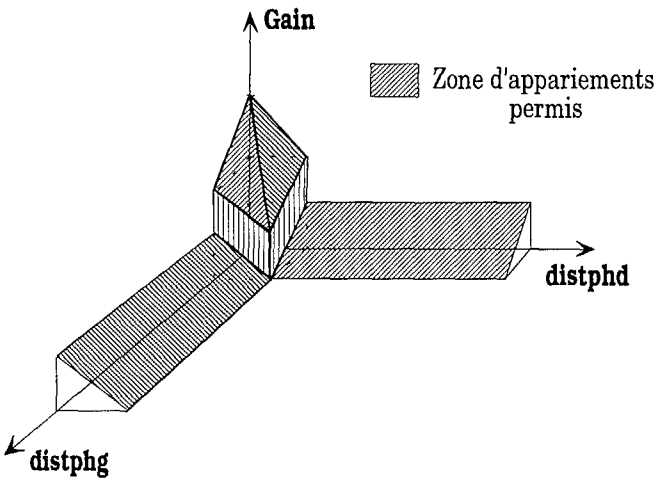


figure 1 : Variation du gain en fonction des distances photométriques

3. Résultats expérimentaux et conclusions

Notre algorithme a été testé sur plusieurs images réelles numérisées au format 512 x 512 x 8 bits. Les résultats obtenus prouvent sa rapidité et sa robustesse. En effet, l'application de notre méthode à des scènes d'intérieur et d'extérieur donne des résultats satisfaisants.

En ce qui concerne le temps de traitement, la mise en correspondance de deux lignes, comportant chacune une trentaine d'éléments caractéristiques, ne demande que 10 ms sur un ordinateur compatible AT 486 cadencé à 33 Mhz, soit environ 5 s pour 2 images complètes, et ceci avec un taux d'appariement correct dépassant 95%.

Nous présentons ci-dessous une carte de profondeur relative à une scène d'intérieur. Cette carte est obtenue par interpolation des résultats établis par notre algorithme de mise en correspondance. Le niveau de gris y est d'autant plus clair que la profondeur est faible.

Signalons pour terminer que l'ensemble du processus de stéréovision que nous venons de présenter est actuellement en cours d'implémentation dans une architecture spécialisée comportant des DSP et des ASIC. Celle-ci exploite largement le parallélisme des algorithmes ce qui permettra de réduire le temps total d'établissement d'une carte de profondeur à un dixième de seconde.

D/G	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	153			30																					
2					151		37																		
3						114																			
4								158																	
5									149																
6										129			38												
7											131	19			39										
8											21	152				16									
9													125	19			35								
10															90	105									
11													46	124			12	35							
12															145				137	23					
13																141				31		33			
14																	133				29		43		
15																		19	32			47		23	132
16																38				32	137		49		
17																	23	45				140			44
18																				34	97		123		
19																							98		43
20																									137
21																								24	

figure 2 : Chemin optimal pour la mise en correspondance de deux lignes épipolaires

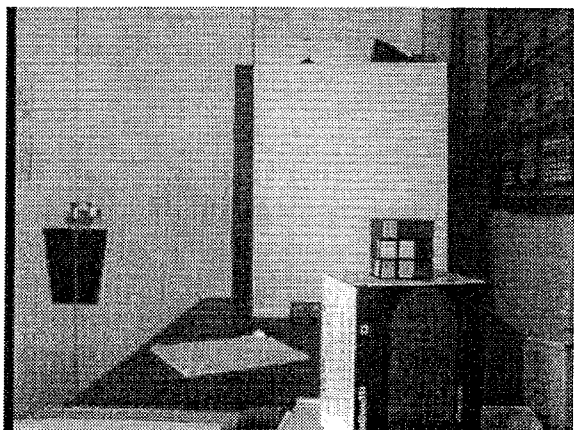


IMAGE GAUCHE

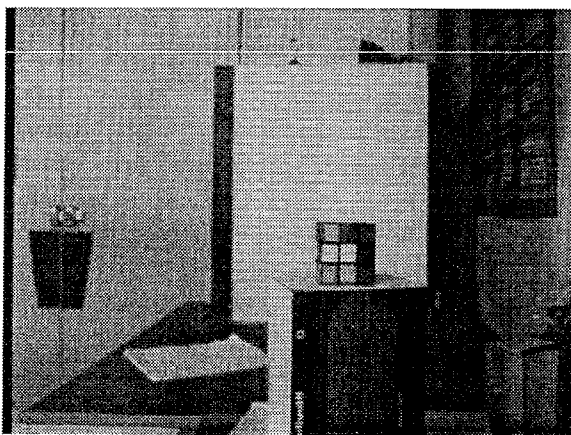
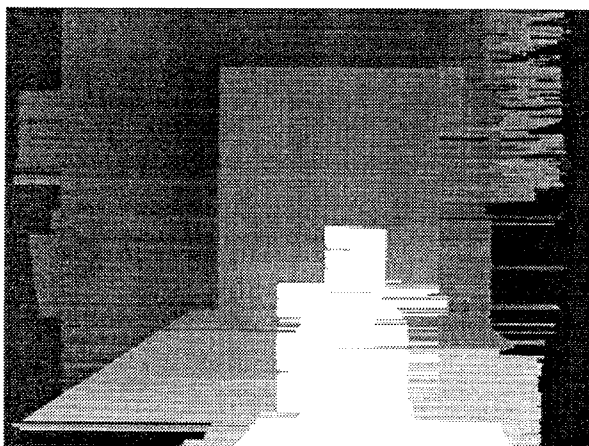


IMAGE DROITE



CARTE DE PROFONDEUR

Références bibliographiques :

- [1] H. H. Baker and T. O. Binford "Depth from edge and intensity based stereo" *Proc. 7th International Joint Conference on Artificial Intelligence*, 631-636, Vancouver Canada , August 1981.
- [2] M. Benard, "Automatic Stereophotogrammetry: A Method Based on Feature Detection and Dynamic Programming", *Photogrammetria*, vol.39, pp. 169-181, 1984.
- [3] Y. Ohta and T. Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming " *IEEE PAMI* 7(2), 139-154, March, 1985.
- [4] S. A. Lloyd, E. R. Haddow and S. E. Dreyfus, "A Parallel Binocular Stereo Algorithm Utilising Dynamic Programming and Relaxation Labelling", *Computer Vision Graphics And Image Processing*, 39, 202-225, 1987.
- [5] Y. Wu and H. Maître, "Mise en correspondance d'images stéréo par programmation dynamique utilisant la cohérence inter lignes " *12^{ème} Colloque GRETSI* Juan-Les-pins Juin 89.
- [6] X. W. Tu and B. Dubuisson, "3-D Information Derivation from a Pair of Binocular Images" *Pattern Recognition*, vol 23, n° 3/4, 223-235, 1990.
- [7] T. Quiguer, P. Miché, R. Debrie "Méthode d'extraction de primitives par seuillage automatique" *13^{ème} Colloque GRETSI* Juan-Les-pins Septembre 91.
- [8] A. Bensrhair, P. Miché and R. Debrie " New Fast Stereo Matching For Parallel Processing" *Proc. IEEE International Conference on Advanced Motion Control* Nagoya, Japan Mars 1992.