



MISE EN CORRESPONDANCE D'IMAGES STEREO PAR PROGRAMMATION DYNAMIQUE UTILISANT LA COHERENCE INTER LIGNES

Yifeng WU - Henri MAÎTRE

Département Images
École Nationale Supérieure des Télécommunications
46, Rue Barrault, 75 013 PARIS

RÉSUMÉ

Un algorithme de Programmation Dynamique est proposé pour la mise en correspondance d'images stéréo. L'algorithme est basé sur une recherche de correspondance intra-lignes épipolaires, et la cohérence inter-lignes est assurée par utilisation de la continuité de disparité le long des contours. L'avantage de cette approche est de nous permettre de garantir la cohérence inter-ligne sans recherche combinatoire dans un espace tri-dimensionnel. Une étape de corrélation est effectuée pour corriger les erreurs de mise en correspondance, elle est suivie par une interpolation, elle est aussi faite par programmation dynamique, après un test d'existence de parties cachées.

SUMMARY

A Dynamic Programming algorithm is proposed in this paper for matching stereoscopic images. The approach is based on an intra-line search at the pixel level. To maintain the continuity of the object outside epipolar lines, the gradient of the contour is used to give a strong constraint for inter-line coherence. The advantage of this method is to ensure the inter-line coherence without the combinatorial search in 3D space. A correlation step is then performed to correct matching errors, and an interpolation is done, using also dynamic programming, after testing if occlusions exist.

I. INTRODUCTION

Le problème de mise en correspondance est le problème le plus difficile de la vision stéréoscopique. La difficulté provient du fait qu'il existe toujours des déformations (déformation géométrique, déformation photométrique, bruit, etc.) entre les deux images. Pour le résoudre plusieurs approches ont été explorées dans la littérature, parmi lesquelles deux conviennent particulièrement bien à la mise en correspondance pixel à pixel : la Programmation Dynamique (PD) et la Corrélation croisée.

La PD est un exemple d'algorithme "coopératif", elle se base sur le principe d'optimalité [BELL-62] : les associations locales de primitives conditionnent la recherche d'une optimisation globale. Cette méthode admet non seulement des primitives de bas niveau comme le pixel, mais aussi des primitives modérément structurées.

Dans les littératures, deux types d'approches de PD sont proposés pour la mise en correspondance d'images stéréo. La première approche [BAKE-81, BENA-83] recherche les correspondances entre une paire de lignes épipolaires conjuguées (intra-ligne), ce qui consiste à minimiser une fonction de coût dans un graphe bi-dimensionnel, et l'algorithme se répète ligne par ligne pour

compléter toute l'image. La deuxième approche [OHTA-85], [LLOY-87] recherche un chemin optimal non seulement en intra-ligne mais aussi en inter-lignes, ce qui consiste à faire une recherche optimale dans un espace 3D. Cette dernière méthode prend bien en compte la cohérence inter-lignes, et donne une solution meilleure que la première. Le défaut majeur de cette approche est sa gourmandise en temps-machine.

Nous proposerons ici une troisième méthode utilisant d'une autre façon l'information inter-lignes.

Cette méthode est basée sur une recherche intra-ligne de correspondance entre les deux images selon la méthode de Viterbi [BAKE-81]. Un gradient directionnel est d'abord calculé sur l'image originale en chaque pixel, qui donne à la fois le module du gradient et son orientation. On introduit ces informations dans la fonction de coût de la PD ; on peut donc garantir la continuité du contour (sur les contours verticaux ou obliques) et contrôler la pente de la variation (sur les contours horizontaux). Puis le résultat donné par la PD est repris par une étape de correction d'erreur utilisant la corrélation [MORA-79, GOSH-84]. Cette correction intègre bien l'information de ressemblance locale; lorsqu'un premier recalage est fait, elle peut fonctionner assez rapidement. Toutes ces performances correspondent bien à notre exigence de correction de l'erreur d'appariement.



Enfin une étape d'interpolation est effectuée par PD. L'existence des parties cachées influence fortement le résultat de l'interpolation. Nous proposons ici de détecter d'abord, l'existence des parties cachées, en testant des fonctions de corrélation entre deux demi-fenêtres de chaque côté du contour; nous pouvons ensuite décider si le point de départ (ou le point final) de PD est fixé ou non, ce qui nous assure un contour net dans la carte de disparité.

II. MISE EN CORRESPONDANCE PAR PD

A partir de deux images redressées géométriquement, la mise en correspondance est effectuée sur chaque paire de lignes épipolaires conjuguées pixel à pixel. Le choix de la fonction de coût joue un rôle très important pour la qualité de l'algorithme. Dans notre cas, le coût élémentaire (coût d'un pas) est composé de 4 parties :

- le coût de ressemblance : CRE,
- le coût de variation : CVA,
- le coût de non correspondance : CNC,
- le coût de cohérence : CCO.

On désigne par $IMG(i)$ le signal d'image gauche (pris pour référence) à l'étape i de la PD, et par $IMD(u(i)), (u(i)=j)$ le signal de l'image droite sur la même ligne épipolaire qui lui est associé à cette étape i .

Le coût de ressemblance représente le changement de niveaux de gris entre deux couples de points successifs le long d'un même chemin :

$$CRE(i-1, u(i-1), i, j) = ABS[(IMG(i) - IMD(j)) - (IMG(i-1) - IMD(u(i-1)))]$$

Le coût de variation représente la ressemblance de la pente du signal d'image entre les deux images :

$$CVA(i, j) = ABS[(IMG(i+1) - IMG(i-1)) - (IMD(j+1) - IMD(j-1))]$$

Le coût de non-correspondance pénalise l'effet de non-correspondance, lorsque une primitive d'une image ne trouve pas de point homologue dans l'autre image, c'est à dire lorsqu'un chemin non diagonal (horizontal ou vertical) est utilisé dans le graphe :

$$CNC = \begin{cases} 0 & \text{si le chemin est diagonal} \\ k & \text{si le chemin est non-diagonal} \end{cases}$$

On a choisi $k=5$ après quelques essais.

Nous introduisons un nouveau coût, le coût de cohérence, pour garantir la continuité de disparité en inter-lignes. Il est évident que le long d'un contour, la disparité doit être continue, c'est-à-dire qu'entre lignes voisines, la différence de disparité doit être faible le long d'un contour. Dans le cas où le contour est horizontal, la variation de la disparité en inter ligne doit être proportionnelle au module du gradient. Pour cela, on calcule d'abord le module et la direction du gradient pour chaque point, et ensuite on définit un seuil de maximum de différence de la disparité en inter-ligne :

$$MAXDIF = \begin{cases} 1 & \text{le long du contour} \\ GRAD/k_1 & \text{perpendiculairement au contour} \end{cases}$$

où $GRAD$ est le module du gradient, k_1 est une constante (on choisit $k_1=10$ dans ce travail). Si la différence de disparité est supérieure à ce seuil, il faut pénaliser cette association :

$$CCO = DIF - MAXDIF$$

où DIF représente la différence réelle.

Après les définitions des différents coûts, nous pouvons écrire le coût élémentaire :

$$c(i-1, u(i-1), i, j) = CRE(i-1, u(i-1), i, j) + CVA(i, j) + CNC + CCO$$

et le coût cumulé :

$$C(i, j) = C(i-1, u(i-1)) + c(i-1, u(i-1), i, j)$$

La minimisation est opérée de gauche à droite le long d'une ligne de l'image gauche; on retient le chemin dont le coût cumulé est minimal, qui représente la meilleure correspondance entre les deux lignes. L'algorithme se répète ligne par ligne pour toute l'image.

III. ALGORITHME DE CORRÉLATION

La PD nous fournit un premier recalage entre les deux images, mais il existe quand même certaines erreurs car la déformation entre les deux images est parfois très grande, il est donc difficile de trouver un bon compromis en réglant la fonction de coût. Nous utilisons donc une étape de correction par corrélation.

La corrélation croisée des images est définie par :

$$C(l, k) = \frac{\sum_{i, j \in w} [IMG(i, j) - M_g][IMD(i+l, j+k) - M_d]}{\sigma_g \sigma_d}$$

où w est une fenêtre de taille fixe (5×5). $M_g, M_d, \sigma_g, \sigma_d$ sont les moyennes et les écarts-type des fenêtres dans chaque image. En déplaçant la fenêtre dans une zone définie, nous déterminons la meilleure correspondance par :

$$C_{\max} = \underset{l, k}{\text{Max}} C(l, k)$$

Comme nous sommes en géométrie épipolaire, k reste constant; on fait varier simplement l autour du résultat donné par la PD. La corrélation est effectuée uniquement sur les points de contour, les seuls points où elle est significative [MORA-79]. On choisit un seuil de corrélation égal à 0,9; malgré ce choix assez sévère, il y a environ 80 % des points testés qui vérifient ce critère. Ceci nous permet de garantir que toutes les disparités sur les contours sont correctes. Une statistique a été faite pour tous les points de contour corrigés par la corrélation, montrant qu'un point sur deux a subi une correction d'un pixel, ce qui confirme l'importance de cette étape de corrélation.

IV. INTERPOLATION

Quelle que soit la méthode utilisée, le processus de mise en correspondance n'est pas réalisé sur tous les points

et dans tous les cas la carte des disparités obtenue n'est pas complète. Cependant la perception visuelle est clairement une surface complète interpolant une représentation initiale incomplète.

L'interpolation est faite ligne par ligne sur chaque intervalle à interpoler, c'est à dire entre deux contours. On juge d'abord s'il y a des parties cachées à ses deux extrémités par un test des fonctions de corrélation entre deux demi-fenêtre de chaque côté de ces points extrémités. Lorsqu'une partie cachée se présente, un écart important apparaît entre elles.

Deux méthodes différentes sont adoptées selon qu'il y a ou qu'il n'y a pas de parties cachées:

- une interpolation linéaire est utilisée s'il n'y a aucune partie cachée aux deux extrémités ou si l'intervalle est assez court (<5 pixels par exemple).
- une interpolation par PD est adoptée lorsqu'une partie cachée se présente. Nous savons que s'il existe une partie cachée la profondeur ne sera pas continue de part et d'autre du contour. Il faut donc laisser ce côté libre quand on utilise la PD: on ne force donc pas, au point de départ (ou au point final), la disparité à être continue. Il faut remarquer que si des parties cachées existent sur les deux côtés, on laisse seulement flotter le côté dont la différence de corrélation est la plus forte car la PD ne permet pas de trouver un chemin satisfaisant sur une zone quasi-uniforme si l'on ne connaît ni le point de départ ni le point d'arrivée. C'est actuellement l'une des faiblesses de l'algorithme.

V. RÉSULTATS

Deux paires d'images stéréo ont été testées par l'algorithme proposé ci-dessus. La figure 1 est issue de la base de test de l'Université de Stuttgart. La taille de ces images est 240×240. Dans ces images, la variation de profondeur est relativement continue, et l'effet de partie cachée n'est pas très fort, par contre, le changement de photométrie est très important entre les deux images. L'image de disparité est présentée en figure 1.(c), et l'image de relief est montrée en figure 1.(d).

La figure 2 montre un exemple d'un BATIMENT, la taille des images originales est 512×512. Sur ces images, l'effet de partie cachée est très important. La figure 2.(c) illustre l'image des disparités, et montre que les parties cachées sont bien détectées, et qui nous donne des changements de profondeur très marqués.

Le temps de calcul de tout l'algorithme est beaucoup plus rapide que la recherche dans un graphe 3D [OHTA-85]. Il faut 2,5 minutes de temps CPU pour les images du mur, et 10 minutes pour les images du bâtiment, sur un Micro Vax-II.

V. CONCLUSION ET DISCUSSION

La combinaison de la PD et de la corrélation nous fournit un outil fiable et rapide pour la stéréorestitution. La PD tient bien compte de l'information globale. Par l'introduction d'un coût de cohérence inter-ligne, et en recherchant uniquement dans un graphe 2D, elle peut nous

donner un premier recalage assez correct dans un temps de calcul très rapide. La corrélation intègre bien l'information locale, et donne une mise en correspondance précise sur les points de contours. Puisque le résultat de PD donne une contrainte géométrique très sévère, le domaine de recherche de la corrélation est largement réduit. Ceci accélère fortement la mise en correspondance.

D'un point de vue général, le résultat du calcul du relief est assez bon. Cependant, il existe encore des erreurs de détails dues essentiellement aux défauts de détection de contour. En effet, dans les régions de surfaces planes, aucune information ne permet de faire l'appariement. Donc théoriquement, si l'on n'est pas capable de faire la segmentation, la reconstitution de la surface ne sera jamais parfaite. Malgré les très gros efforts faits en détection de contours, le résultat obtenu n'est pas encore totalement satisfaisant. Récemment, une nouvelle méthode de détection de contours est apparue [BERG-87, SJOB-88] basée sur la distinction des contours et des ombrages. C'est une voie probablement intéressante pour améliorer la qualité du contour.

BIBLIOGRAPHIE

- [BAKE-81] H.H.Baker and T.O.Binford, "Depth from edge and intensity based stereo", Proc. 7th International Joint Conference on Artificial Intelligence, pp.631-636, Vancouver, British Columbia, August, 1981.
- [BELL-62] R.E.Bellman and S.E.Dreyfus, "Applied Dynamic Programming", Princeton University Press, Princeton, N.J., 1962.
- [BENA-84] M.Bénard, "Automatic Stereophotogrammetry: A Method Based on Feature Detection and Dynamic Programming", Photogrammetria, Vol.39, pp.169-181, 1984.
- [BERG-87] F.Bergholm, "Edge Focusing" IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.PAMI-9, No.6, pp.726-741, 1987.
- [GOSH-84] A.Goshtasby, S.H.Gage and J.F.Bartholic, "A two-stage cross correlation approach to template matching", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.PAMI-6, No.3, pp.374-378, May, 1984.
- [LLOY-87] S.A.Lloyd, E.R.Haddow and S.E.Dreyfus, "A Parallel Binocular Stereo Algorithm Utilizing Dynamic Programming and Relaxation Labelling", Computer Vision, Graphics, and Image Processing, 39, pp. 202-225, 1987.
- [MORA-79] H.P.Moravec, "Towards Automatic Visual Obstacle Avoidance", Proc. 5th IJCAI, pp.584, Cambridge, MA, Aug. 1977.
- [OHTA-85] Y.Ohta and T.Kanade, "Stereo by intra- and inter-scanline search using dynamic programming", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.PAMI-7, No.2, pp.139-154, March, 1985.
- [SJOB-88] F.Sjoberg and F.Bergholm, "Extraction of Diffuse Edges by Edge Focusing", Pattern Recognition Letters, 7, pp.181-190, March, 1988.

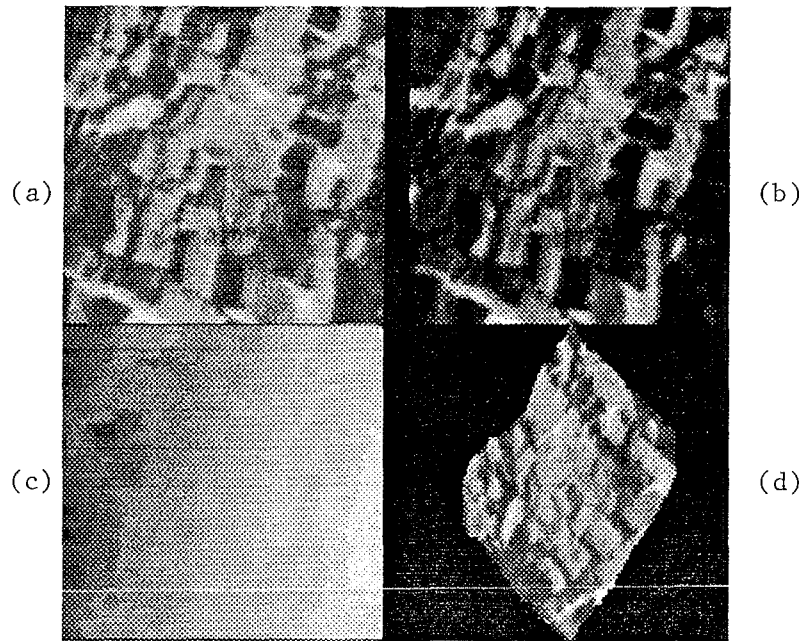


Figure 1: MUR
 (a) image gauche, (b) image droite, (c) la carte de disparité, (d) vue perspective.

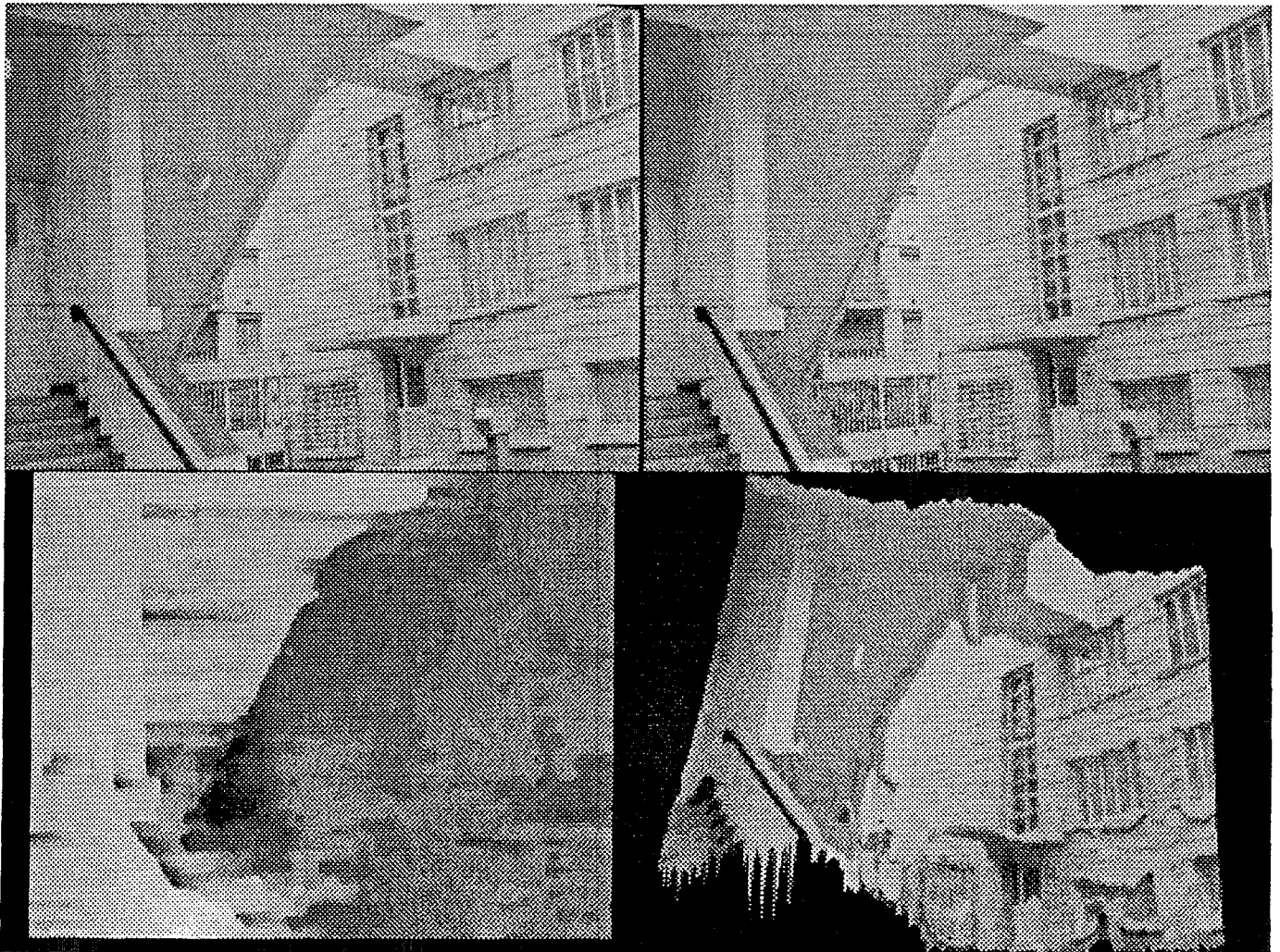


Figure 2: BATIMENT
 (a) image gauche, (b) image droite, (c) la carte de disparité, (d) vue perspective.