

Un Algorithme Multi-Echelle de Mise en Correspondance Stéréo Basé sur les Champs Markoviens

Antonin Chambolle¹ Davi Geiger²
Stéphane Mallat³

¹CEREMADE, Univ. de Paris-Dauphine, Paris ²Siemens Corp. Res., Princeton, NJ
³Robotics Lab., Courant Institute of Math. Sciences, New York University, New York

Résumé Nous présentons un algorithme de mise en correspondance stéréo par une méthode basée sur les champs de Markov, qui intègre la phase d'interpolation de la disparité et de détection de ses discontinuités et la mise en correspondance proprement dite d'éléments caractéristiques en une seule étape. Ces éléments sont obtenus par transformée en ondelettes des images, à des échelles d'abord grossières puis de plus en plus fines.

1 Introduction : la Stéréo

Le principe de la vision stéréoscopique est a priori simple : il s'agit d'exploiter l'information de relief existant en théorie dans une paire de vues d'une même scène prises sous des angles légèrement différents. Les problèmes fondamentaux à résoudre sont celui de la connaissance de la géométrie et de la calibration du système de prise de vues, et celui de la mise en correspondance des points des deux images. On va supposer le premier résolu : la paire d'images est alors telle que les points à associer sont sur la même ligne $y = \text{constante}$ (*contrainte épipolaire*) et la *disparité* $d(x, y)$ qui mesure le déplacement entre deux de ces points le long d'une ligne épipolaire est positive (ou négative) ou nulle sur l'ensemble de l'image ($d = 0$ correspondant aux points placés à l'infini). Grâce à cette contrainte la mise en correspondance peut se faire ligne à ligne et le problème est essentiellement unidimensionnel.

Pour le résoudre, on peut essayer d'associer soit les points d'égal intensité lumineuse (R. March 1988), en tenant compte éventuellement des fluctuations entre les deux images (Gennert 1988), soit plutôt des figures caractéristiques de la scène ("image features") en général plus stables et plus fiables : coins, bords (Grimson 1981, Marr 1982, [3]).

Quelle que soit la méthode utilisée on n'obtient jamais d'information complète, en particulier la deuxième approche ne peut produire qu'un résultat très parsemé. Un problème ultérieur est de remplir les trous, ce qui se fait en établissant des hypothèses a priori sur la régularité des surfaces observées, qui se traduisent par certaines hypothèses sur d : d est continue par morceaux, dérivable par morceaux, on lui suppose aussi une forte régularité aux endroits où on ne dispose d'aucune information. (De telles hypothèses sont naturelles puisque les parties les plus homogènes de l'image correspondent en général à des surfaces régulières.) Une autre hypothèse naturelle est de supposer pour tout y la fonction $x \rightarrow d(x, y)$ croissante ("monotonicity constraint") : c'est-à-dire qu'un objet situé à gauche d'un autre dans une des deux images ne peut être à sa droite dans l'autre. Ce n'est pas tou-

Abstract We introduce a stereomatching algorithm based on the Markov Random Field approach to vision. It solves simultaneously the interpolation of the disparity function (and the detection of its breaks) and the matching of image features. Those features are extracted from the wavelet transforms of the images, at coarse and then finer space scales.

jours vrai dans la réalité, mais il semble que même l'œil admette une hypothèse semblable, et ne traite pas avec succès les situations où ce principe est violé. Le fait que d ne soit continue que par morceaux est lié à un problème important : à y fixé si, au point x , $d(x^-, y) \neq d(x^+, y)$ alors tous les points de la deuxième image situés entre $x + d(x^-, y)$ et $x + d(x^+, y)$ n'ont pas de correspondants dans la première image (en vertu du principe de monotonie), et appartiennent à une *zone d'occlusion*. Symétriquement il peut aussi exister de telles zones dans la première image, sur lesquelles la disparité n'est pas vraiment définie. C'est un problème très délicat qui pourtant ne présente pour l'œil aucune difficulté.

2 Stéréo et physique statistique

L'apparition de méthodes de physique statistique en traitement d'images est récente (D. & S. Geman 1984, Azencott 1987, [5]). L'idée principale consiste à considérer l'image comme un réseau de particules (les points) qui peuvent se trouver dans différents états (en général en nombre fini). A chaque état est associée une certaine énergie, avec des termes d'interactions entre particules voisines, de telle sorte que l'ensemble puisse être analysé comme un champ markovien.

En stéréo, on propose une approche de ce type (voir [5]), dont voici un bref exposé en dimension un :

Les signaux ont N points d'abscisses $x_n = n.s$ où s est un pas d'échelle, $n \in [0, N-1]$, et contiennent chacun des figures caractéristiques aux points x_{l_1}, \dots, x_{l_M} dans le premier et en y_{r_1}, \dots, y_{r_P} dans le second. A chaque paire (x_{l_i}, y_{r_j}) est associé un coefficient ϕ_{l_i, r_j} mesurant la ressemblance entre la forme du premier signal en x_{l_i} et du second en y_{r_j} : ϕ_{l_i, r_j} est d'autant plus petit que les signaux sont plus semblables, et donc que x_{l_i} doit être associé à y_{r_j} avec plus de certitude.

On note d_l la disparité (inconnue) au point x_{l_i} et δ_{l_i, r_j} la disparité en x_{l_i} si x_{l_i} est associé à y_{r_j} (qui n'est pas nécessairement $y_{r_j} - x_{l_i} = (r_j - l_i)s$ car en réalité on a souvent une information plus précise sur les abscisses réelles des figures caractéristiques — qui seront des extrema de transformée en



ondelettes, voir la section suivante — situées en x_{l_i} et y_{l_i} , que les valeurs arrondies x_{l_i} et y_{l_i} qui proviennent d'un sous-échantillonnage.)

On va chercher à minimiser l'énergie $E = \sum_{l=0}^{N-1} E_l$, où

$$\begin{cases} E_l = \sum_{j=1}^P M_{l_i, r_j} \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\} + E_l^r \text{ si } l = l_i \\ E_l = E_l^r \text{ si } l \neq l_i \forall i \\ \text{où } E_l^r = \alpha(d_l - d_{l-1})^2(1 - h_l) + \gamma h_l \text{ (} l \geq 1 \text{)} \end{cases} \quad (1)$$

$$\text{Avec : } M_{l_i, r_j} = \begin{cases} 1 & \text{si } l_i \text{ et } r_j \text{ sont associés} \\ 0 & \text{sinon ;} \end{cases} \quad (2)$$

(il est clair qu'une contrainte devra exprimer le fait qu'à chaque l_i ne peut correspondre qu'un r_j au plus, et réciproquement).

h_l est le processus de discontinuité plus souvent appelé processus de ligne (horizontal) (ou "line process") (en deux dimensions) : $h_l = 1$ si d peut être discontinue entre x_{l-1} et x_l , $h_l = 0$ sinon et dans ce cas le terme $\alpha(d_l - d_{l-1})^2$ de la partie régularisante de l'énergie E^r est actif.

α, γ sont des paramètres positifs, qu'il faudra estimer, ils permettent de régler le poids relatif et les valeurs critiques du processus de régularisation. Lorsque $\gamma = 0$ on remarque par exemple que E^r est nulle et minimale quand $\forall l, h_l = 1$, et rien n'oblige la solution à être régulière.

En vue de minimiser l'énergie E on commence d'abord par essayer d'éliminer les variables les plus gênantes, et c'est là qu'intervient la physique statistique. Pour une température inverse β fixée la probabilité de se trouver dans un état donné d'énergie E est $e^{-\beta E}/Z$, où Z est la fonction de partition :

$$Z(\beta) = \sum_{\{\text{tous les états du système}\}} e^{-\beta E} = \sum_{\{d\}} \left(\sum_{\{h, M\}} e^{-\beta E} \right) = \sum_{\{d\}} Z^d(\beta, d) \quad (3)$$

La sommation sur tous les états définis par les variables h_l est facile, en revanche, pour sommer sur les différents états de M_{l_i, r_j} on est amené à faire une hypothèse qui éloigne un peu le modèle des objectifs : on suppose que chaque élément caractéristique x_{l_i} de la première image s'associe à un et un seul élément y_{r_j} de la deuxième, sans autre contrainte ; en particulier on n'empêche pas a priori un y_{r_j} d'être associé à deux x_{l_i} et $x_{l_i'}$ différents. Cette hypothèse se traduit par la condition

$$\sum_{j=1}^P M_{l_i, r_j} = 1 ; \text{ et il vient :} \quad (4)$$

$$Z^d = \prod_{i=1}^M \left(\sum_{j=1}^P e^{-\beta \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\}} \right) \prod_{l=1}^{N-1} \left(e^{-\beta \alpha (d_l - d_{l-1})^2} + e^{-\beta \gamma} \right).$$

On calcule alors l'énergie effective $E_{\text{eff}}(\beta, d) = -\frac{1}{\beta} \ln(Z^d)$ qu'on va essayer de minimiser par des méthodes de gradient descendant. Cette minimisation permet d'intégrer en une seule procédure la mise en correspondance des éléments caractéristiques et la régularisation de d entre ces éléments, alors que la plupart des autres algorithmes procèdent par étapes. A haute température (β petit) E_{eff} est presque convexe en d , et on peut facilement approcher son minimum. Ensuite, en faisant croître β on "refroidit" le système en espérant pouvoir suivre le minimum jusqu'à un état proche de $\beta = +\infty$ (température nulle), où il correspond effectivement à l'état le plus probable du système, soit celui où l'énergie (1) est minimale. Cette démarche est à rapprocher de la méthode dite *G.N.C* (Graduate Non Convexity) exposée par A. Blake et

A. Zisserman (1987). On peut calculer les valeurs moyennes des champs h_l et M_{l_i, r_j} :

$$\begin{aligned} \bar{h}_l &= \frac{1}{1 + e^{-\beta \alpha (d_l - d_{l-1})^2 - \gamma}} \\ \bar{M}_{l_i, r_j} &= \frac{e^{-\beta \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\}}}{\sum_{j'=1}^P e^{-\beta \{(d_{l_i} - \delta_{l_i, r_{j'}})^2 + \phi_{l_i, r_{j'}}\}}} \end{aligned} \quad (5)$$

et on voit qu'à température nulle les conditions (2), (4) et ($h_l \in \{0, 1\}$) sont réalisées.

On comprend bien le rôle des paramètres α, γ dans l'expression (5) : si $|d_l - d_{l-1}| < \sqrt{\frac{\gamma}{\alpha}}$ on a $h_l \approx 0$ et d doit être continue entre x_{l-1} et x_l , alors que si $|d_l - d_{l-1}| > \sqrt{\frac{\gamma}{\alpha}}$, $h_l \approx 1$ et la force attirant d_l et d_{l-1} l'un vers l'autre se relâche : $\sqrt{\frac{\gamma}{\alpha}}$ est un seuil de gradient pour d (ce qui se dégage également des formules (1)).

En fait, les deux signaux stéréo ne sont pas identiques ni même translatés mais plutôt déformés l'un par rapport à l'autre et il arrive qu'un élément x_{l_i} ne corresponde à aucun y_{r_j} . On peut pallier cet inconvénient en introduisant un seuil ϵ au dessus duquel il est plus avantageux de ne pas associer x_{l_i} à un des y_{r_j} : pratiquement, cela se fait en remplaçant la condition (4) par :

$$\sum_{j=1}^P M_{l_i, r_j} \in \{0, 1\} \quad (6)$$

et, pour rejeter la solution triviale $M_{l_i, r_j} = 0 \forall i, j$, en rajoutant un terme la pénalisant dans l'énergie qui devient si $l = l_i$: $E_l =$

$$\sum_{j=1}^P M_{l_i, r_j} \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\} + \epsilon(1 - \sum_{j=1}^P M_{l_i, r_j}) + E_l^r.$$

L'énergie effective devient à une constante près $E_{\text{eff}}(\beta) =$

$$\begin{aligned} & -\frac{1}{\beta} \left\{ \sum_{i=1}^M \ln \left(1 + \sum_{j=1}^P e^{-\beta \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\} - \epsilon} \right) \right. \\ & \left. + \sum_{l=1}^{N-1} \ln \left(e^{-\beta [\alpha (d_l - d_{l-1})^2 - \gamma]} + 1 \right) \right\}, \text{ et :} \\ \bar{M}_{l_i, r_j} &= \frac{e^{-\beta \{(d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j}\} - \epsilon}}{1 + \sum_{j'=1}^P e^{-\beta \{(d_{l_i} - \delta_{l_i, r_{j'}})^2 + \phi_{l_i, r_{j'}}\} - \epsilon}} \end{aligned} \quad (7)$$

Le problème des zones d'occlusion est traité comme suit : on suppose, au début, d définie partout, et on effectue d'abord une mise en correspondance de l'image de gauche vers celle de droite (on aura remarqué que la méthode exposée a le défaut de ne pas traiter symétriquement les deux images, puisque la fonction de disparité n'est définie que sur l'une, et aussi du fait de l'hypothèse traduite par la condition (4)) ; on obtient une fonction $d^{L \rightarrow R}$ qui est une certaine approximation de la fonction de disparité définie sur l'image de gauche. On calcule alors à partir de cette fonction une approximation de la disparité $d^{R \rightarrow L}$, en utilisant la relation :

$$d^{R \rightarrow L}(x^L + d^{L \rightarrow R}(x^L)) = -d^{L \rightarrow R}(x^L) \quad (8)$$

(Cette transformation n'est pas du tout immédiate parce que les coordonnées x^L et x^R sont entières ou multiples d'un certain pas s alors que $d^{L \rightarrow R}$ est une fonction réelle.) Lorsque la fonction calculée $d^{L \rightarrow R}$ présente une discontinuité entre deux points x_l^L et x_{l+1}^L (c'est-à-dire lorsque le "line process"

h_{l+1} vaut 1), il est impossible d'associer une valeur $d^{R \rightarrow L}$ aux points x_r^R situés entre $x_l^L + d_l^{L \rightarrow R}$ et $x_{l+1}^L + d_{l+1}^{L \rightarrow R}$, qui appartiennent à une zone d'occlusion.

En deux dimensions, le modèle est identique, si ce n'est que la partie régularisante de l'énergie est légèrement différente, puisqu'elle tient compte de la continuité verticale de la disparité, et on doit introduire des variables de processus de ligne vertical $v_{i,j}$.

3 Primitives Utilisées pour la Mise en Correspondance.

Dans la réalisation d'un système de mise en correspondance une étape fondamentale est de déterminer les primitives que l'on va chercher à associer d'une image à l'autre. L'idée la plus élémentaire consiste à faire se correspondre des points d'égale intensité lumineuse, mais cette information est instable, les surfaces réelles n'étant généralement pas lambertiennes et aussi du fait des fluctuations et erreurs introduites par les systèmes de saisie des images. L'information réellement significative en stéréo est surtout contenue dans les hautes et moyennes fréquences horizontales des images, et en particulier les bords — discontinuités importantes de l'intensité. Ces bords sont aussi les lieux où les discontinuités physiques de la scène réelle ont le plus de chance de se produire, puisqu'elles s'accompagnent en général d'une brusque variation d'intensité lumineuse : la recherche des discontinuités de la disparité et des zones d'occlusion s'aidera donc avec profit de la connaissance des bords. La représentation bas niveau utilisée par notre programme, basée sur la transformée en ondelettes [2], est essentiellement équivalente aux primitives extraites au moyen d'un détecteur de bords de J. Canny (Canny 1986). Elle offre les avantages suivants : elle présente un caractère multi-échelle¹ bien maîtrisé dans le cadre de la théorie des ondelettes dyadiques ; elle est presque complète et contient toute l'information initialement présente dans la paire d'images stéréo (voir [1]).

$\Theta(x, y)$ étant, une fonction régularisante, on introduit deux ondelettes directionnelles : une ondelette "horizontale" $\Psi^1(x, y)$ et une "verticale" $\Psi^2(x, y)$ où :

$$\Psi^1(x, y) = \frac{\partial \Theta(x, y)}{\partial dx} \quad \text{et} \quad \Psi^2(x, y) = \frac{\partial \Theta(x, y)}{\partial dy} \quad (9)$$

On note $\Psi_{2^j}^1(x, y) = \frac{1}{2^{2j}} \Psi^1(\frac{x}{2^j}, \frac{y}{2^j})$, etc... et la transformée en ondelettes d'une fonction f à l'échelle 2^j est :

$$\begin{aligned} W_{2^j}^1 f(x, y) &= f * \Psi_{2^j}^1(x, y) \quad (\text{composante horizontale}) \\ W_{2^j}^2 f(x, y) &= f * \Psi_{2^j}^2(x, y) \quad (\text{composante verticale}) \end{aligned} \quad (10)$$

L'ondelette est interprétée comme un filtre passe-bande que l'on dilate. A grande échelle j , $W_{2^j}^i f(x)$ ($i = 1, 2$) détecte les basses fréquences du signal f , et lorsque j décroît le support de $\Psi_{2^j}^i$ se réduit et $W_{2^j}^i f(x)$ devient sensible à des détails de plus en plus fins : l'échelle j caractérise la taille et la régularité des singularités du signal analysé par la transformée. On voit que :

$$\begin{cases} W_{2^j}^1 f(x, y) = 2^j \frac{\partial}{\partial x} (f * \Theta_{2^j})(x, y) \\ W_{2^j}^2 f(x, y) = 2^j \frac{\partial}{\partial y} (f * \Theta_{2^j})(x, y) \end{cases}, \quad (11)$$

¹La plupart des chercheurs ont insisté sur le fait qu'un système efficace de mise en correspondance devait nécessairement traiter les signaux à différentes échelles (cf par exemple [4]), en commençant par essayer d'associer les détails grossiers de la paire d'images pour petit à petit pouvoir associer les détails de plus en plus fins en utilisant évidemment l'information obtenue aux étapes précédentes.

ainsi les maxima locaux de $|W_{2^j}^1 f(x, y)|$ le long de la direction x , à y constant, sont les points où $f(x, y)$ régularisée à l'échelle 2^j varie horizontalement le plus rapidement, et l'ensemble de ces points quand y varie forme des courbes correspondant aux bords à peu près verticaux de l'image à l'échelle 2^j . Si on considère les maxima locaux de $|W_{2^j}^2 f(x, y)|$ le long de y on obtient les bords horizontaux de l'image, et si on considère les arêtes de $|W_{2^j}^1 f(x, y)|^2 + |W_{2^j}^2 f(x, y)|^2$ on obtient plus ou moins un détecteur de bords multi-échelle très similaire à celui de J. Canny. Pour la mise en correspondance, les lignes épipolaires étant les lignes ($y = \text{constante}$), on s'intéressera surtout au premier ensemble de courbes (les bords verticaux).

4 Implémentation

Le programme réalisé se décrit comme suit :

- Calcul de la transformée en ondelettes (des fonctions de niveau de gris) des images, sur n niveaux.
- Recherche des maxima correspondant aux bords verticaux, dont la position est interpolée en adaptant une fonction quadratique sur les trois points définissant chaque maximum.

Ensuite, à chaque échelle ($J = n - 1$ à 0) :

- On calcule les quantités ϕ_{i,r_j} où i est l'indice d'un maximum de la première image, en (x_i^1, y_i^1) et j d'un maximum de la deuxième, en (x_r^2, y_r^2) ($y_i^1 = y_r^2 = y$). ϕ_{i,r_j} est calculé en effectuant des corrélations entre les transformées en ondelettes à l'échelle 2^J des images aux voisinages des deux maxima. On effectue un premier seuillage en considérant le couple de maxima (i, j) incompatible si ϕ_{i,r_j} est au-delà d'une certaine valeur.
- On minimise à plusieurs reprises, en diminuant à chaque fois la température, l'énergie effective associée au système à l'échelle J . (Cette énergie a toutefois été légèrement modifiée pour rendre les discontinuités de d beaucoup plus probables aux points où des bords ont été repérés dans l'image.) La condition initiale est la valeur de d calculée aux cycles précédents, ou zéro la première fois. La minimisation suit une méthode de gradient descendant : (ici en dimension un)

$$\begin{aligned} d_l^{m+1} &= d_l^m - \omega \left(\frac{\partial E_{\text{eff}}}{\partial d_l} (d^m) \right) \\ &= d_l^m - 2\omega \left(\alpha (d_l^m - d_{l-1}^m) (1 - \bar{h}_l^m) \right. \\ &\quad \left. + \alpha (d_l^m - d_{l+1}^m) (1 - \bar{h}_{l+1}^m) \right. \\ &\quad \left. + \sum_{j=1}^P \bar{M}_{l_i, r_j}^m (d_l^m - \delta_{l_i, r_j}) \right) \end{aligned} \quad (12)$$

(la dernière ligne n'apparaît que lorsque $l = l_i$) ; elle s'arrête lorsqu'entre deux pas successifs m et $m + 1$ l'énergie a diminué moins qu'un certain seuil.

Remarque Le paramètre β joue entre autres le rôle d'une "fenêtre" comparable à celle utilisée dans [3] : en effet la quantité $e^{-\beta((d_{l_i} - \delta_{l_i, r_j})^2 + \phi_{l_i, r_j})}$ est (presque) nulle dès que $(d_{l_i} - \delta_{l_i, r_j})^2 \gg \frac{1}{\beta}$, par exemple initialement si la disparité recherchée δ_{l_i, r_j} est de l'ordre de 25 pixels et la condition initiale impose $d_{l_i} = 0$, on doit avoir (si on ne veut pas que le programme soit complètement "aveugle") $625 \sim \frac{1}{\beta}$ et donc il vaut mieux commencer la minimisation avec des valeurs de β de l'ordre de 10^{-3} . En augmentant la valeur de β d'étape en étape (en "refroidissant" le système), on réduit de plus en plus le champ



d'exploration de l'énergie effective et du processus de minimisation autour de la valeur initiale de d — sensée être de plus en plus précise et proche de la disparité cherchée — qu'on lui fournit.

4. On renverse la disparité qu'on vient de calculer, en utilisant la formule (8). Les zones d'occlusions de la disparité originale deviennent des discontinuités du résultat, et réciproquement. Les zones d'occlusion sont signalées (par un tableau de booléens) et aucune disparité n'y sera calculée à l'étape suivante.
5. A ce moment, soit on reprend au point 2, soit, lorsqu'on a effectué le calcul de la disparité à l'échelle J dans les deux sens (signal de gauche vers signal de droite et réciproquement), on passe à l'échelle suivante en remplaçant J par $J - 1$ et en interpolant la disparité sur une grille deux fois plus fine (le pas d'échelle s est divisé par deux). On reprend alors au point 2 (tant que $J \geq 0$). (Le point 2 lui-même n'a en fait pas a priori à être répété tant qu'on ne change pas d'échelle J .)

On a constaté qu'à chaque changement d'échelle on pouvait diviser γ par deux : en effet tant que les signaux sont très régularisés, aux échelles grossières, chercher d'éventuelles discontinuités de la disparité n'a pas grand sens, et il vaut mieux que le seuil $\sqrt{\gamma/\alpha}$ soit grand, ce n'est que lors des dernières étapes qu'on cherche à voir les discontinuités apparaître.

5 Résultats

On remarque qu'essentiellement:

- sur les zones continues la mise en correspondance est très bien effectuée et révèle (lorsque α n'est pas trop grand) des détails fins du relief.
- les discontinuités sont généralement bien détectées lorsqu'elles sont significatives, même si leur localisation n'est pas toujours parfaite en particulier du fait de leur tendance à développer des profils "en escalier" au lieu de suivre correctement les contours des objets — ce qui serait le cas si on avait disposé d'une image de contours débarrassée de tout les bords qui ne correspondent pas à des discontinuités réelles de profondeur.

La figure 2 représente une vue tridimensionnelle de la fonction de disparité calculée pour une paire stéréo dont la figure 1 est une des deux images. Dans la figure suivante la disparité est représentée en niveaux de gris, les zones d'occlusion étant à la valeur zéro (gris moyen). Cette vue montre mieux les discontinuités qui ont été repérées

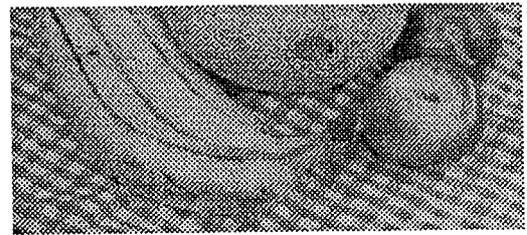


Figure 1: fruits (image de gauche)

The Fruits

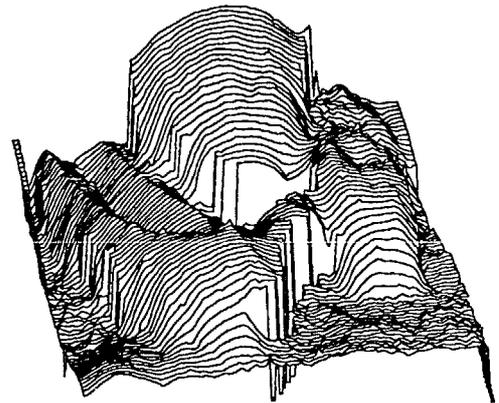


Figure 2: vue tridimensionnelle

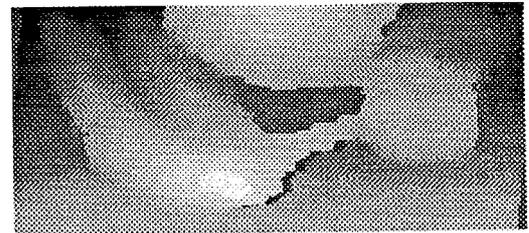


Figure 3: vue en niveaux de gris

Références

- [1] Stéphane G. Mallat and Sifen Zhong. Complete signal representation with multiscale edges (New York University computer science technical report). 1990. To appear in *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [2] Yves Meyer. *Ondelettes et Opérateurs*. Hermann, 1989.
- [3] Soren I. Olsen. Stereo correspondence by surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, March 1990.
- [4] Andrew Witkin, Demetri Terzopoulos, and Michael Kass. Signal matching through scale space. *International Journal of Computer Vision*, 133-144, 1987.
- [5] A. L. Yuille, D. Geiger, and H. Bülthoff. Stereo integration, mean field theory and psychophysics. In *ECCV 90 proceedings, Lecture Notes in Computer Science 427*, pages 73-82, Springer Verlag, April 1990.