

STEREOSCOPIE PAR MISE EN CORRESPONDANCE DE GRAPHES

STEREO VIEWING BY GRAPH MATCHING

Marc BERTHOD - Pascale LONG

I.N.R.I.A. - Route des Lucioles - SOPHIA ANTIPOLIS - 06560 VALBONNE - FRANCE

**RESUME**Résumé:

Nous présentons un algorithme qui extrait partiellement l'information 3D d'une scène à l'aide d'une méthode stéréoscopique.

Nous utilisons une paire d'image obtenue par enregistrement d'une même scène sous deux angles différents.

La méthode utilisée ici est l'amélioration d'étiquetage. Le modèle du monde est défini de façon à satisfaire les contraintes propres à la vision stéréo, telles que les hypothèses d'unicité et de continuité.

Dans notre formalisme, un objet est une correspondance potentielle entre deux points issus chacun d'une des images, une étiquette dénote la présence ou l'absence de cet objet.

Le traitement nécessite en outre, une analyse de l'étiquetage final pour affecter une et une seule étiquette à chaque objet. Nous voulons trouver toutes les correspondances affectées de l'étiquette "présent". De tels objets sont définis par des couples de points qui représentent les projections d'un même point physique.

La correspondance établie, on peut connaître la disparité associée aux objets qui conduit à une notion de profondeur.

L'algorithme a été appliqué sur trois types d'images qui se distinguent par la nature des candidats à la correspondance: les pixels d'un stéréogramme aléatoire, des éléments de contours et des points d'intérêts sélectionnés par un prétraitement sur les images.

**SUMMARY**Abstract

We propose an algorithm which partially retrieve 3D information of a scene by a stereoscopic method. We use a pair of images of the same scene taken from two different viewpoints.

This approach is an application of relaxation-optimization labeling technique which satisfies the stereo problem constraints: unicity and continuity.

According to that, an object is a possible correspondence between two points respectively belonging to each image. A label corresponds to the existence or the non-existence of this object.

The final step is an analysis of the resulting labeling, in order to choose a unique label for each object.

This selection is based on neighbourhood information and the unicity constraints. Our purpose is to find all the objects with an "existing" label. Such an object represents a satisfying match between two points: they could be interpreted as image plane projections of the same real world surface point.

The disparity corresponding to this object, provide depth information.

Three kinds of images have been taken as algorithm input. They differ in their candidats to be matched:

- . pixels of the random dot stereogram
- . edge elements
- . points of interest obtained by a pre-processing.



## 1. Introduction.

La restitution de l'information tridimensionnelle (plus simplement '3D') à partir d'une paire de vues stéréoscopiques d'une scène, suscite actuellement un regain d'intérêt, notamment en raison des applications à la vision en robotique. Le principe de base de la stéréoscopie est de déterminer des couples de pixels correspondant aux projections de mêmes points 3D sur chacun des récepteurs. Un simple calcul de trigonométrie permet alors d'en déduire les coordonnées 3D de ces points.

Si l'on considère la vision stéréoscopique humaine, le problème de la mise en correspondance semble trivial. Malheureusement, comme c'est le plus souvent le cas pour les tâches de perception, la facilité avec laquelle l'oeil (ou le cerveau) réalise cette mise en correspondance ne nous éclaire pas, en tout cas pas suffisamment, sur la manière dont il le fait.

De fait, la réalisation d'un système automatique extrayant l'information 3D à partir de vues stéréoscopiques s'est avérée un problème très difficile.

Le point crucial est qu'il existe en général pour un élément donné d'une vue, disons la vue gauche, de nombreux correspondants possibles sur la vue droite, si l'on considère seulement la ressemblance entre les paramètres associés à ces points (intensité lumineuse dans le cas d'images en niveaux de gris). La correspondance ne peut donc s'établir sans ambiguïté qu'à un niveau plus global, ce qui implique un processus coopératif pour la prise en compte du contexte de chaque point.

De nombreuses manières d'aborder ce problème, fondées ou non sur des tentatives de modélisation de la vision stéréo humaine [1] ont été proposées. Elles se distinguent tout d'abord par le choix des éléments à mettre en correspondance: pixels [2,3], éléments de contours [1], points d'intérêt particulier [4].

Le type d'algorithmes mis en jeu varie aussi: corrélation [3,5], relaxation [4], assignation directe, et combinaison de plusieurs résolutions. [1]

Toutes ces méthodes ont cependant en commun de prendre en compte, implicitement ou explicitement, les particularités de la vision stéréo. Ces particularités ont été résumées par Marr [1] dans des hypothèses classiques.

### 1 Hypothèse d'unicité.

Chaque point d'une vue a un correspondant unique sur l'autre vue. Cette condition repose sur l'hypothèse qu'une correspondance représente un objet physique dont la position dans l'espace est unique.

### 2 Hypothèse de continuité.

A deux points voisins, appartenant à une même surface visible, correspond une faible variation de profondeur.

Ces lois seront violées chaque fois qu'il y aura des discontinuités de profondeur. Cependant, une

petite fraction de l'image, constituée essentiellement par les bords d'objets, est concernée.

Tout le problème est de satisfaire ces règles presque partout parfaitement, et non pas partout presque parfaitement. Il est donc assez naturel de formaliser le problème sous la forme d'une mise en correspondance de graphes. Par exemple, chacune des vues de la paire stéréo est décrite par un graphe. Les sommets d'un graphe sont les candidats à mettre en correspondance; les arêtes schématisent des contraintes. Les correspondants de deux sommets reliés par une arête sont choisis en fonction des deux règles précitées. Puisqu'en certains points, les règles d'unicité et de continuité sont violées, c'est plus précisément des isomorphismes simultanés de sous-graphes qui sont cherchés.

Nous proposons ici d'utiliser l'amélioration d'étiquetage [6,7] pour résoudre le problème de la mise en correspondance de graphes. Une importante particularité du formalisme choisi est que les deux vues jouent un rôle symétrique, ce qu'on obtient par une modification simple du schéma ci-dessus.

## 2. Amélioration d'étiquetage.

La nature du problème à traiter permet de constituer un ensemble  $O$  d'objets (de dimension  $N$ ) et un ensemble  $L$  d'étiquettes (de dimension  $K$ ).

L'étiquetage associe à chaque couple objet-étiquette une mesure  $P_{ik}$  de la confiance dans l'assignation de cette étiquette à cet objet. Dans ce qui suit, on pose  $P = (P_{11}, \dots, P_{1k}, \dots)$ . Diverses contraintes peuvent en outre être imposées à  $P$ . Le plus souvent, les  $P_i$  sont interprétés comme des vecteurs de probabilité; dans ce cas  $\forall i, k, P_{ik} \geq 0$ , et  $\forall i, \sum_k P_{ik} = 1$ .

L'amélioration d'étiquetage consiste à coordonner les choix des étiquettes pour chaque objet. La notion de contexte est donc à la base de cette méthode.

Pour chaque objet  $o_i$ , on considère un ensemble  $V_i$ , appelé voisinage de  $o_i$ , contenant les objets  $o_j$  en relation avec  $o_i$ .

On définit alors un vecteur support  $R$ , appelé vecteur de compatibilité, qui pour chaque couple  $(u_i, l_k)$  fournit une estimation de la fiabilité de l'affectation en fonction du contexte agissant sur lui. C'est aussi une prédiction, sur l'objet en question, de l'influence du contexte qui apparaît sous forme de paramètres. Une méthode simple pour combiner l'influence des différents objets du voisinage, peut se faire sous la forme suivante:

$$R_{ik} = \sum_{j \in V_i} \sum_{l \in L} P_{ik}^0 \cdot C_{ikjl} \cdot P_{jl}$$

$C_{ikjl}$  quantifie l'influence exercée par l'affectation de l'étiquette  $l_1$  à l'objet  $o_j$  de  $V_i$  sur l'affectation de  $l_k$  à  $o_i$ .

Plus  $P_{ik}^0$  et  $R_{ik}$  sont semblables, plus les observations sont compatibles avec le modèle du monde.

Un critère de consistance mesure simultanément la similarité entre les  $P_{ik}$  et  $R_{ik}$ , en tenant compte des contraintes imposées à  $P$ . On a choisi comme critère le produit scalaire entre  $P$  et  $R$ .

$P^*$  est un étiquetage consistant si:

$$\begin{cases} C(P^*) = (P^{*T} \cdot R) \max \\ \sum_{k \in GL} P_{ik} = 1, P_{ik} > 0 \quad \forall i, k \end{cases}$$

Il est en général impossible d'appliquer à ce système des méthodes quadratiques, compte tenu du nombre de contraintes imposées. De plus, les méthodes de gradient ne permettent d'obtenir qu'un optimum local, dépendant fortement du choix de l'étiquetage initial.

Par contre, si l'on admet une normalisation des  $R_{ik}$  pour chaque objet  $u_i$  on peut modifier la définition de la consistance pour obtenir une solution au problème. Cette modification s'appuie sur l'interprétation de  $\lambda_i$  comme étant la probabilité a priori d'occurrence des observations pour l'objet  $u_i$

$$\forall i, k \begin{cases} \sum_{k \in GL} R_{ik} = \lambda_i, \lambda_i > 0 \\ \sum_{k \in GL} P_{ik} = 1, P_{ik} > 0 \end{cases}$$

et la semi-consistance est définie par:

$$R_{ik} = \sum_{j \in V_1} \sum_{l \in L} P_{ik}^0 \cdot C_{ikjl} \cdot (\lambda_j / \lambda_i) \cdot P_{jl}$$

On considère la grandeur  $X_{ik}$ , telle que  $X_{ik} = \lambda_i \cdot P_{ik}$  comme la composante  $(i, k)$  d'un vecteur  $X$  et  $(P_{ik}^0 \cdot C_{ikjl})$  comme la composante  $M_{ijkl}$  d'une matrice  $M$ . L'équation peut s'exprimer sous forme matricielle:

$$\lambda \cdot R = M \cdot X$$

Par l'intermédiaire de  $X^*$ , on recherche une solution au problème qui maximise le critère de semi-consistance entre l'étiquetage et le modèle du monde.

$$\begin{cases} C(X^*) = (X^{*T} \cdot M \cdot X^*) \max \\ \|X\|_2 = 1 \end{cases}$$

Ce vecteur  $X^*$  est donc un extremum de  $C(X)$ , sur la sphère de rayon unité dans l'espace de dimension  $N \times K$ .

Si  $M$  est non négative et irréductible, il existe un vecteur solution unique  $X^*$ , strictement positif. C'est le vecteur propre associé à la plus grande valeur propre de  $(M + M^T)$ .

Ces propriétés de la matrice  $M$  sont souvent vérifiées car  $M$  est non négative et son graphe associé est fortement connexe.

### 3. Réduction à un problème à deux étiquettes.

Nous avons choisi de traiter le problème de correspondance stéréoscopique comme une recherche d'isomorphisme de sous-graphes.

Comme nous l'avons cité précédemment, chacune des vues de la paire stéréo est décrite sous forme de graphe.

Pour un graphe donné, les sommets sont les éléments de l'image candidats pour la mise en correspondance;

les arêtes représentent des contraintes liant deux points voisins.

Pour réaliser la mise en correspondance de sous-graphes par une méthode d'amélioration d'étiquetage, nous devons définir un formalisme adapté à la nature du problème.

Par exemple, l'objet  $u_i$  peut être un élément d'une des vues, l'étiquette  $l_k$  un élément similaire dans l'autre image. Le couple  $(u_i, l_k)$  constitue alors un couple susceptible d'être une correspondance entre les deux images. Dans le cas d'une interprétation probabiliste,  $P_{ik}$  représente la probabilité pour l'objet  $u_i$  d'avoir l'étiquette  $l_k$ .

Nous pouvons remarquer que les deux graphes jouent, dans ce cas, le même rôle: c'est-à-dire qu'il y a dualité entre les objets et les étiquettes.

Pour respecter cette symétrie, et prendre en compte la possibilité d'avoir un nombre variable d'étiquettes pour un objet, nous proposons de prendre comme objets les couples objet-étiquette  $(u_i, l_k)$  et comme étiquettes les étiquettes "présent" et "absent". Donc, à un objet  $O_{ik}$  sont associées les grandeurs  $P_{ik}$  (affectation de l'étiquette  $l_k$  dans le formalisme de base), et  $A_{ik}$  (non affectation de cette étiquette).

Dans le cas de l'isomorphisme de graphes, un objet devient donc un couple candidat à une correspondance.

Dans le cas plus particulier de la vision stéréo, un objet  $O_{ik}$  est un point 3D potentiel.  $P_{ik}$  s'interprète comme une mesure de la qualité de cette correspondance,  $A_{ik}$  comme une mesure de l'action pénalisante de l'objet  $O_{ik}$

### 4. Interprétation de l'étiquetage final

Le vecteur  $X^*$  résultat est donc le vecteur propre associé à la plus grande valeur propre de la matrice  $(M + M^T)$ .

$X^*$  contient tous les  $P_{ik}$  et  $A_{ik}$  du nouveau formalisme:

$$\begin{aligned} X_{ik,p}^* &= \lambda_{ik} \cdot P_{ik}^* \\ X_{ik,a}^* &= \lambda_{ik} \cdot A_{ik}^* \end{aligned}$$

Les éléments de  $X^*$  sont étudiés dans le but de sélectionner les objets pour qui l'étiquette "présent" est retenue, puisque c'est le problème qui nous intéresse.

En effet, nous pourrions alors dire que le couple d'éléments lui correspondant est constitué par les projections d'un point 3D sur chacune des images.

Cette sélection est réalisée en deux étapes:

A. La première est l'estimation d'une grandeur  $S(O_{ik})$  définie pour chaque objet par:

$$S(O_{ik}) = P_{ik}^* - A_{ik}^*$$

$S(O_{ik})$  est fonction croissante de  $X_{ik,p}$  et décroissante de  $X_{ik,a}$ . Les valeurs de  $S(O_{ik})$  sont comprises entre -1 et 1.



B. La seconde étape consiste à utiliser  $S(O_{ik})$  pour obtenir un tri final non ambigu. Les étiquettes sont incompatibles, et donc on cherche à associer à un objet une et une seule étiquette. Or,  $S(O_{ik})$  apporte une information qui n'est pas directement utilisable. En effet, si  $S(O_{ik})$  est très voisin de 1 (respectivement -1), il y a de fortes raisons d'affecter à  $O_{ik}$  l'étiquette "présent"(absent). Cependant, pour la plupart des objets,  $S(O_{ik})$  ne permet pas de faire un choix fiable pour une étiquette plutôt qu'une autre. Ceci exclut toute sélection par un seuil sur les valeurs de  $S(O_{ik})$ . La valeur du seuil serait très délicate à déterminer, et nous ne pourrions pas éviter les problèmes d'ambiguïtés.

Nous avons choisi une méthode de sélection qui tient compte des contraintes d'unicité définies au paragraphe 1.

Un objet  $O_{nm}$  est formé d'une origine  $a_n$  dans l'image de gauche associée à une extrémité  $b_m$  dans l'image de droite,  $O_{nm} = (a_n, b_m)$ .

$$\text{Soit } E1_n = [O_{nk} / \neg b_k, O_{nk} = (a_n, b_k)]$$

$E1_n$  regroupe tous les objets d'origine  $a_n$ . De même  $E2_m$  est l'ensemble des objets d'extrémité  $b_m$ .

$$E2_m = [O_{*m} / \neg a_k, O_{*m} = (a_k, b_m)]$$

Avec ce formalisme, si  $O_{nk} = (a_n, b_k) \in E1_n$  est tel que  $S(O_{nk})$  est maximum, on peut dire que  $b_k$  est l'extrémité la plus probable pour l'origine  $a_n$ .

Donc, on peut définir un test de sélection sous la forme suivante:

$$\text{Si } \neg O_{nm} \text{ tel que } \begin{cases} O_{nm} \in E1_n / S(O_{nm}) \text{ max} \\ O_{nm} \in E2_m / S(O_{nm}) \text{ max} \end{cases}$$

alors, l'étiquette "présent" est affecté à  $O_{nm}$ , et l'étiquette "absent" à tous les autres objets appartenant à  $E1_n$  et à  $E2_m$ .

Ces objets, ayant leur étiquette finale, ne sont plus considérés dans la suite du traitement.

Si, après avoir étudié tous les cas, il n'existe pas de  $O_{ik}$  vérifiant les deux conditions simultanément, on recommence la même opération en réduisant l'ensemble des objets à ceux pour lesquels l'étiquette finale n'est pas déterminée.

C'est un traitement itératif, et on associe aux objets un coefficient qui mesure la fiabilité de l'affectation de l'étiquette. Ce coefficient dépend du nombre d'itérations nécessaire pour déterminer l'étiquette finale de l'objet. Après un nombre fini d'itérations, les objets non étiquetés seront considérés comme des situations impossibles. C'est le cas des points 3D qui ne se projettent que dans une des images (parties cachées, bords).

##### 5. Prétraitement sur les images.

Il est clair qu'un tel traitement ne peut pas s'appliquer à tous les pixels d'une image car le nombre d'objets  $O_{ik}$  deviendrait trop important. Pour

éviter ce problème, nous sélectionnons dans les images seulement certains points d'intérêt.

Les opérateurs utilisés doivent permettre de distinguer clairement un point de ses voisins. Ils peuvent être sensibles aux variances locales, ou à d'autres propriétés très particulières qui serviront à assurer une grande fiabilité sur la position des points détectés.

Nous avons choisi pour le prétraitement des images l'opérateur de Moravec. [8] Les points sont sélectionnés de la manière suivante:

On cherche, pour chaque point, le minimum parmi les quatre directions (horizontale, verticale, et les deux diagonales) de la variance directionnelle, c'est à dire la somme des carrés de la différence entre les valeurs des pixels sur une fenêtre  $5 \times 5$ .

Cette valeur,  $V(i,j)$ , est grande seulement si toutes les variances directionnelles sont élevées.

Le point  $(i,j)$  est retenu si cette grandeur est maximale localement, et si elle est supérieure à un seuil fixé pour s'assurer de la qualité du choix. Ce maximum local est défini sur les huit plus proches voisins du point.

##### 6. Etudes des résultats

Au cours de la mise en place de l'algorithme, nous avons dû remédier à deux problèmes spécifiques à la stéréo:

- la recherche de ligne épipolaire pour limiter le nombre d'objets du matching
- la différence d'apparence entre les deux images

Le problème de recherche de la direction épipolaire est résolu si la calibration des caméras, et leur position relative sont connues. De plus, ces paramètres sont nécessaires pour exploiter, par des considérations optiques et géométriques, les résultats de notre traitement.

Il est important de limiter l'ensemble des correspondants possibles pour un point, tout en s'assurant que le bon candidat soit contenu dans cette liste. Pour un point P de l'image M1, nous avons choisi de chercher les correspondants parmi les points de l'image M2 contenus dans une bande. Cette bande est orientée suivant la direction épipolaire, elle est centrée autour du point ayant les mêmes coordonnées que P dans M2. Sa longueur dépend du domaine de profondeur autorisé.

Les images d'une paire stéréo sont prises sous deux angles de vues différents. Il est donc fréquent que les conditions d'éclairage ou de réflectance ne soient pas les mêmes. Ces phénomènes provoquent des variations de "qualité" entre les images (par exemple pour les clichés aériens).

En conséquence, la similitude entre les niveaux de gris des candidats à une correspondance ne peut pas être un critère discriminant. Nous utilisons des paramètres dépendant des propriétés locales dans l'image. Par exemple, la cohérence entre leur voisinage direct apporte une information beaucoup plus fiable. De cette manière, l'algorithme est moins sensible aux variations de "qualité" d'une image par rap-



port à l'autre.

On appelle disparité, l'écart relatif des coordonnées entre les deux éléments en correspondance. C'est cette grandeur qui permet d'aboutir à une notion de profondeur. Pour la vision humaine, le fond de la scène correspond à une disparité nulle.

L'algorithme a été testé dans un premier temps, sur des stéréogrammes aléatoires binaires. Les candidats à la correspondance sont alors les pixels. Il est, en effet, nécessaire de travailler au plus bas niveau car il n'y a aucune relation dans l'image permettant de constituer des éléments plus précis. Cela permet de mettre en valeur l'importance de l'information contextuelle dans un traitement de ce type. Dans ce cas, elle n'est fonction que de la proximité des objets de même disparité. Par analogie à la vision humaine, nous n'avons considéré que des écarts latéraux (ligne épipolaire horizontale).

Compte tenu de ces remarques, l'application de l'algorithme permet rapidement de situer les zones de même profondeur. Aux frontières de ces régions, on peut constater quelques problèmes d'ambiguïté dus au manque d'information sur l'environnement des objets lorsqu'il y a des variations brutales de disparité.

Pour étudier le problème de différence de qualité entre les deux vues, nous avons testé l'algorithme sur un stéréogramme du même type dont une des images est bruitée suivant une distribution gaussienne. Les résultats correspondent assez bien à ceux obtenus précédemment. Les problèmes d'ambiguïté se ressentent toujours, mais il est clair que le manque d'information est amplifié sur une image bruitée. Si ce n'est pas le cas, seul le hasard au cours de la distribution du bruit en est responsable.

Le manque d'information pour réaliser la sélection des candidats, permet de prévoir que les résultats s'amélioreront si les images ne sont pas binaires. En référence à la remarque sur la qualité d'image, la notion de niveaux de gris influe peu sur la création des objets. Par contre, elle intervient pour différencier un couple Oik des couples voisins. En effet, la valeur moyenne d'un couple (demi somme des pixels) est peu sensible aux variations de qualité des images si les dégradations sont localement homogènes. Les contraintes sont donc plus fortes.

Une modification se ressent dans l'application sur ce second type de stéréogramme. Les valeurs des pixels ont été distribuées aléatoirement suivant une loi uniforme. Nous avons noté une amélioration dans les zones délicates tel que les coins. Le nombre d'itérations nécessaires pour obtenir un résultat satisfaisant est inférieur au cas des stéréogrammes binaires. Donc la convergence est plus rapide.

Dans un deuxième temps, cet algorithme a été appliqué à des images dont on a extrait les éléments de contours.

Le nombre de candidats est considérablement réduit. Le détecteur de contours a pour effet de sélectionner dans l'image des points ayant la propriété particulière d'appartenir à un bord d'objet.

Cependant, les objets de la scène, projetés sur des surfaces planes, ont des contours déformés dans une image par rapport à l'autre. De tels éléments ne se

distinguent pas suffisamment de leurs voisins pour aboutir à un étiquetage consistant. Pour ces raisons, les points de contours sont mal appropriés à une application de ce type.

C'est la raison pour laquelle nous avons utilisé l'opérateur de Moravec, dans la phase de prétraitement. Cet opérateur permet d'extraire des points d'intérêts dans les images. Les points sélectionnés vérifient des propriétés particulières qui permettent d'avoir une connaissance précise de leur position. L'ensemble des candidats se limite donc aux points significatifs de l'image. Nous obtenons entre 100 et 500 points, sur une image 128x128, suivant la complexité de la scène étudiée. De plus, l'information qu'ils transportent facilite leur identification pour l'étape de recherche des correspondants.

Les points étant espacés, nous pouvons agrandir la taille du voisinage renforceur sans augmenter considérablement les temps de calcul. Ceci permet de rajouter des contraintes sur l'objet étudié, et d'accélérer la convergence.

L'opérateur est peu sensible aux différences de qualité des images: les points M1 et M2, projections dans chacunes des vues d'un même point M de la scène 3D, vérifient des propriétés qui dépendent peu des conditions d'éclaircissement.

Donc, dans le cas idéal, les points extraits dans chaque image correspondent aux mêmes objets de la scène réelle. Quelques parties de cette scène n'apparaissent que sur une des vues de la paire stéréo. Cela est dû soit à la limite du champ visible dans les deux caméras, soit aux effets produits par les parties cachées et les ombres.

Les résultats tendent à montrer que la plus grande majorité des correspondances réalisées sont exactes. Par des calculs géométriques, nous pourrions trouver les profondeurs de ces points 3D par rapport aux caméras.

## 7. Conclusion

Le vecteur support R, reflétant l'influence du contexte, est une forme quadratique linéaire qui fait intervenir les voisins de l'objet considéré.

$$R_{ik} = \sum_{j \in V_1} \sum_{l \in L} P_{ik}^0 \cdot C_{ikjl} \cdot P_{jl}$$

Si on fait une interprétation probabiliste de R, on peut noter que les voisins interviennent de façon indépendante, et surtout que leur effet n'est pas simultané.

D'une part, l'hypothèse d'indépendance est souvent abusive: on suppose que les voisins  $j_1$  et  $j_2$  de  $V_1$  sont indépendants, alors qu'il y a de grandes chances pour que  $j_1$  appartienne au voisinage de  $j_2$  et inversement. Mais cette hypothèse ne provoque pas de gros bouleversements et simplifie beaucoup les calculs.

D'autre part, il s'agit d'une somme sur les voisins, qui correspond à un "ou" probabiliste. Il serait plus réaliste de décrire l'action globale de tous les voisins, agissant avec une étiquette ou une autre.

Une telle conception du problème est représentée par une expression légèrement différente de celle utilisée ici:



$$R_{ik} = \prod_{j \in V_1} \sum_{l \in L} P_{ik}^o \cdot C_{ikjl} \cdot P_{jl}$$

Cette nouvelle forme n'est plus linéaire, et peut poser certains problèmes à la mise en oeuvre, mais il serait intéressant de comparer les résultats des deux méthodes.

Nous sommes en train de travailler sur une méthode plus fine pour déterminer le choix des correspondants. Cette modification permettrait d'appliquer l'algorithme plus généralement à la mise en correspondance d'image.

#### References

1. Marr D. and Poggio T., A Theory of Human Stereo Vision, November 1977.
2. Julesz B., "Towards the Automation of Binocular Depth Perception," in Proceedings of the IFIPS congress, Munich (1962).
3. Hannah M. J., "Computer Matching of Areas in Stereo Images," Stanford A.I Vol. memo 239 (July 1974).
4. Barnard S.T. and Thompson W.B., "Disparity Analysis of Images," IEEE Trans. on Pattern Analysis and Machine Intelligence Vol. PAMI-2, NO 4, pp.333-340 (July 1980 ).
5. Yakimovsky Y. and Cunningham R., "A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras," Computer Graphics and Image Processing Vol. 7, pp.195-210 (1978).
6. Faugeras O.D. and Berthod M., "Improving Consistency and Reducing Ambiguity in Stochastic Labeling: an Optimization Approach," IEEE Trans on PAMI Vol. 3(4), pp.412-423 (1981).
7. Berthod M. and Faugeras O.D., "Using Context in the Global Recognition of a Set of Objects: an Optimization Approach," Proc. of the 8th W.C.C. (IFIP), p.695 (1980).
8. Moravec H.P., "Towards automatic visual obstacle avoidance ," pp. 584 in Proc. 5th Int. Joint Conf. Artificial Intell., Cambridge (Aug. 1977).