

Rectification d'images et Reconstruction 3-D par Analyse des Régions en Stéréovision

Ezzeddine Zagrouba, Charlie Krey
ENSEEIH - IRIT - UA 1399

Laboratoire "Vision par ordinateur A. Bruel"
2 Rue Camichel, 31071 TOULOUSE Cedex - FRANCE
Tél : (33) 61 58 83 79 ; Fax : (33) 61 58 82 09 ;
e-mail : zagrouba@enseeiht.fr / krey@enseeiht.fr

Résumé

Ce papier présente une méthode de reconstruction 3-D à partir de la mise en correspondance des régions d'une paire d'images stéréoscopiques. L'originalité de cette méthode réside dans l'utilisation de primitives 2-D riches en informations de surface (régions) et des contraintes 2-D issues de l'hypothèse de planéité en 3-D faite sur les scènes traitées. La reconstruction des surfaces tridimensionnelles se fait directement par recalage des frontières des régions, et en effectuant uniquement des calculs sur des primitives 2-D. Nous proposons pour cela un modèle de rectification d'images qui nous permet d'avoir une géométrie simplifiée du banc stéréoscopique. Nous exprimons ensuite les différentes contraintes 2-D dans ce nouveau modèle. Enfin, nous proposons une modélisation du problème de la reconstruction 3-D à partir d'un couple de régions homologues sous forme d'une minimisation d'une fonction énergie. La résolution du problème est faite par un algorithme de recuit simulé.

1 Introduction

Le problème de la reconstruction 3-D en stéréovision a été souvent étudié en utilisant un modèle de surface, en l'occurrence les surfaces planes. Nous distinguons deux familles de techniques de reconstruction des facettes 3-D planes (une surface plane d'un objet opaque de la scène observée) :

- recherche des surfaces (souvent planes) directement dans l'espace 3-D à partir d'une carte de disparité issue d'une analyse des contours ou des points de contraste (utilisation des points 3-D ou des segments 3-D) [2], [5], [7], [13] et [9] ;
- génération des hypothèses dans l'espace 2-D à partir d'un modèle de surface et validation par la recherche du plan correspondant à la carte de disparité obtenue [12] et [3].

Mais, le processus de triangulation, utilisé pour le calcul des coordonnées 3-D à partir des coordonnées images gauche et droite, ne peut qu'accentuer les erreurs cumulées. Celles-ci sont dues aux phénomènes successifs suivants :

- discrétisation du plan focal des caméras ;
- segmentation des images [4] et [8] ;
- calibrage des caméras (paramètres extrinsèques et intrinsèques) [1] et [10]. De plus, les informations fournies par le constructeur sont souvent trop imprécises (facteur d'échelle et distance focale).

L'utilisation d'un tel processus produit une incertitude de reconstruction 3-D non isotrope et fonction de la position du point considéré [3]. Par conséquent, la gestion de l'erreur, nécessaire dans un tel processus, devient une tâche très complexe.

Pour cette raison, nous avons opté pour l'expression des contraintes 3-D en terme de contraintes 2-D dans les plans images. Nous effectuons la reconstruction 3-D en ne travaillant qu'avec des primitives 2-D, où l'hypothèse d'incertitude isotrope sur la détermination des coordonnées des points 2-D est plus réaliste.

2 Notre approche

Le choix de la primitive 2-D "région" a été motivé par la richesse d'une telle primitive en information de surface. En effet, la région, vue comme un polygone 2-D définissant une zone d'intensité homogène, peut être considérée comme une *localisation floue d'une facette 2-D* (projection d'une facette 3-D sur le plan image). Notons que la limitation majeure de cette approche réside dans l'instabilité de la primitive 2-D "région". En effet, les outils de segmentation en régions fournissent des frontières mal localisées à cause d'une grande sensibilité aux variations d'intensité dans les images. Pour répondre à cette exigence nous proposons, d'une part une coopération entre la segmentation et la mise en correspondance [14], [15] et d'autre part, une correction de la disparité entre les frontières des régions homologues. Cette correction est guidée par la contrainte de planéité en 2-D combinée avec une contrainte radiométrique (corrélation des images des niveaux de gris).

Nous nous basons sur la puissance de l'hypothèse de planéité en 3-D que nous transposons en 2-D sous la forme d'une transformation homographique τ_{gd} (dans le cas d'un banc stéréoscopique à géométrie rectifiée) entre les coordonnées du point image gauche (u_g, v_g) et son homologue droit (u_d, v_d). Puis, nous utilisons cette transformation pour corriger les frontières de régions telle que la disparité soit continue à l'intérieur d'une région associée à une facette plane 3-D.

Pour ce faire, nous avons besoin de connaître la matrice de passage entre chacun des deux repères (gauche et droit) du plan image et le repère lié à la caméra correspondante ainsi que la matrice de passage reliant les deux repères liés aux caméras. C'est pourquoi, nous commençons par effectuer une rectification des images suivant un modèle d'un banc stéréoscopique à géométrie rectifiée. Puis, nous explicitons l'expression

de la transformation homographique τ_{gd} en fonction des paramètres de ce nouveau modèle déduit du modèle initial (deux matrices perspectives P_G et P_D).

En effet, nous supposons que l'incertitude sur la localisation d'un point de frontière est une zone de longueur $2\Delta n$ pixels, centrée sur ce point. Nous déterminons le meilleur estimateur de la transformation τ_{gd} associée à chaque couple de régions homologues, en faisant varier les points de frontière à l'intérieur de cette zone d'incertitude (supposée isotrope et non fonction de la position du point dans l'image).

Pour l'implantation d'une telle méthode, nous l'avons modélisé sous forme d'un problème d'optimisation que nous résolvons avec un algorithme de recuit simulé pour éviter les minima locaux.

Le schéma de l'algorithme se résume comme suit :

- | |
|--|
| <p>i) Rectifier les images suivant un modèle à géométrie simple.</p> <p>ii) Estimer les nouveaux paramètres du modèle géométrique.</p> <p>iii) Pour chaque couple des régions homologues faire</p> <ul style="list-style-type: none"> - Initialisation : solution initiale + fonction énergie - Résolution du problème de minimisation : recuit simulé - Estimation de la transformation homographique: $\tau_{gd} = \begin{pmatrix} \alpha_u & \alpha_v & 1 \\ 0 & 1 & 0 \end{pmatrix}$ <ul style="list-style-type: none"> - Calcul de l'équation du plan infini dans (G) puis dans (M) : (p,q,r) et (A,B,C,D) - Limitation de la facette 3-D dans l'espace. |
|--|

3 Modèle de la rectification

L'absence de procédure de calibrage facilement utilisable ajoutée à certains avantages mentionnés précédemment, a amené la plupart des chercheurs en stéréovision à supposer dans leurs algorithmes une géométrie épipolaire simple avec les droites épipolaires parallèles aux lignes image. Or, une telle configuration peut être obtenue par deux méthodes:

- alignement mécanique très précis, ce qui est difficilement réalisable en l'absence d'instrument de mesure à la hauteur de la précision exigée ;
- transformation épipolaire (rectification) pour rendre les droites épipolaires parallèles aux lignes image.

En plus de cette configuration nous avons besoin de connaître les différents paramètres du banc stéréoscopique rectifié. Nous effectuons la rectification d'images suivant un modèle précis (deux caméras dans le même plan, axes optiques parallèles et où le passage du repère lié à la caméra gauche à celui lié à la caméra droite se réduit à une simple translation). Par conséquent, nous sommes en mesure de calculer les nouvelles matrices perspectives M_G et M_D à partir des matrices initiales (P_G et P_D) qui nous permettent :

- d'accélérer le calcul des épipolaires ;
- de calculer la position et l'orientation des deux caméras en plus de certains paramètres internes [10].

D'où l'obtention de la matrice de passage du repère lié à la caméra au repère du plan image correspondant et de la matrice de passage du repère lié à la caméra gauche au repère lié à la caméra droite.

3.1 Notations

Considérons les différentes notations suivantes :

- P_G et P_D : les deux matrices perspectives gauche et droite initiales.
- M_G et M_D : les deux nouvelles matrices perspectives gauche et droite après rectification.
- $(M) = (O, X, Y, Z)$: le repère absolu lié au robot.
- $(G) = (O_G, X_G, Y_G, Z_G)$ [resp. $(D) = (O_D, X_D, Y_D, Z_D)$] : repère lié à la caméra gauche [resp. droite] après rectification.
- $(L) = (o_g, u_g, v_g)$ [resp. $(R) = (o_d, u_d, v_d)$] : repère du plan image gauche [resp. droit] après rectification.
- C_G et C_D : les deux centres optiques des caméras.
- (\mathcal{R}, T) : la rotation et la translation qui permettent de passer de (G) à (M). Soit :
 $\mathcal{R} = (r_1 \ r_2 \ r_3)^t$, $T = (t_x \ t_y \ t_z)^t$
 et $\forall i \in \{1, 2, 3\} : r_i = (r_{i1} \ r_{i2} \ r_{i3})$
- $T_{GD} = (\Delta_x \ 0 \ 0)^t$: vecteur translation pour passer de (G) à (D).

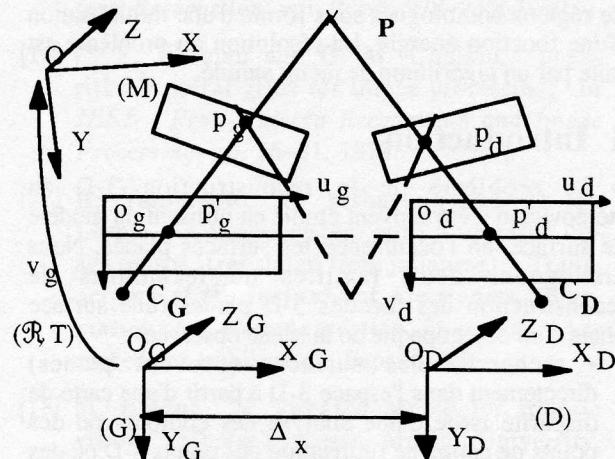


Figure 1 : Modèle de rectification

3.2 Modèle de formation d'image

Le modèle idéal de formation d'image est celui de la projection perspective (sténopé). Pour tenir compte de la non perpendicularité des axes du plan image (formée sur la rétine), [10] propose un modèle de formation de l'image avec correction de la perpendicularité où l'axe des abscisses est parallèle aux lignes image. Un tel modèle permet d'éviter l'utilisation des informations fournies par le constructeur (paramètres intrinsèques des caméras). En effet, ces informations sont entachées d'erreurs. Soit :

$$\begin{pmatrix} X_G & Y_G & Z_G \end{pmatrix}^t = M_{GL} \cdot (s.u_g \ s.v_g \ s)^t \text{ et } \begin{pmatrix} X_D & Y_D & Z_D \end{pmatrix}^t = M_{DR} \cdot (s.u_d \ s.v_d \ s)^t ;$$

$$\text{où : } M_{GL} = M_{DR} = \begin{pmatrix} b & c & a \\ 0 & e & d \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

3.3 Passage de (G) et (D) à (M)

Le passage de (G) à (M) est un déplacement décomposable en une rotation \mathcal{R} et une translation T (voir fig. 1). Soit :

$$\begin{cases} (X_G \ Y_G \ Z_G)^t = \mathcal{R} \cdot (X \ Y \ Z)^t + T \\ (X_D \ Y_D \ Z_D)^t = \mathcal{R} \cdot (X \ Y \ Z)^t + T + T_{GD} \end{cases} \quad (2)$$

3.4 Passage de (G) à (D)

Nous supposons aussi que les deux caméras sont dans le même plan 3-D et que les origines O_G et O_D de (G) et (D) sont situées sur la même horizontale. Ainsi, le passage de (G) à (D) se réduit à la translation T_{GD} suivant l'axe (O_G, X_G) . Soit :

$$(X_D \ Y_D \ Z_D)^t = (X_G \ Y_G \ Z_G - \Delta_x)^t \quad (3)$$

3.5 Nouvelles matrices perspectives

Par combinaison des équations (1), (2) et (3) nous obtenons :

$$\begin{aligned} (s.u_g \ s.v_g \ s)^t &= M_G \cdot (X \ Y \ Z)^t \\ \text{et } (s.u_d \ s.v_d \ s)^t &= M_D \cdot (X \ Y \ Z)^t \end{aligned} \quad (4)$$

$$\text{avec } M_G = \begin{pmatrix} l_1 & l_{14} \\ l_2 & l_{24} \\ l_3 & l_{34} \end{pmatrix} \quad (5)$$

$$\text{et } M_D = \begin{pmatrix} l_1 & l_{14} - \Delta_x \\ l_2 & l_{24} \\ l_3 & l_{34} \end{pmatrix}$$

$$\text{et } \forall i \in \{1, 2, 3\} \ l_i = (l_{i1} \ l_{i2} \ l_{i3})$$

4 Détermination du nouveau modèle

4.1 Calcul de M_G et M_D

Ayache [2] a déjà proposé une méthode de rectification mais sans imposer de contraintes supplémentaires relatives aux paramètres du banc stéréoscopique. Ceci lui a permis de fixer certains degrés de liberté qui restaient disponibles de façon à faciliter la résolution.

Dans notre cas, nous imposons aux nouvelles matrices perspectives M_G et M_D d'être sous la forme (5). Remarquons tout d'abord que la modélisation décrite précédemment nous garantit que :

- les deux plans focaux sont confondus. Donc, les épipolaires sont parallèles aux lignes image ;
- les droites épipolaires conjuguées sont telles que : $v_g = v_d = \text{Cte}$.

De plus, ce modèle nous permet d'obtenir la matrice de passage entre les deux repères (G) et (D). Et par rapport aux trois contraintes définies dans [2], il nous en reste une seule à vérifier, à savoir :

- les centres optiques associés à M_G et M_D doivent être respectivement C_G et C_D et ceci pour assurer

une correspondance univoque entre les points image (p_g, p_d) avant rectification et (p'_g, p'_d) après rectification (voir fig. 2), c'est-à-dire :

$$M_G \cdot \begin{pmatrix} C_G \\ 1 \end{pmatrix} = M_D \cdot \begin{pmatrix} C_D \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

D'où le système de 6 équations à 13 inconnues suivant :

$$\begin{cases} l_1 \cdot C_G + l_{14} = 0 \\ l_1 \cdot C_D + l_{14} = \Delta_x \\ l_2 \cdot C_G + l_{24} = 0 \\ l_2 \cdot C_D + l_{24} = 0 \\ l_3 \cdot C_G + l_{34} = 0 \\ l_3 \cdot C_D + l_{34} = 0 \end{cases}$$

Sachant que les matrices perspectives sont définies à un facteur d'échelle près, nous posons :

$l_{34} = \| C_G \wedge C_D \|^2$. Ce qui nous laisse 6 degrés de liberté disponibles correspondant à la position de (G) par rapport à (M).

Prenons : $l_{24} = 0$; ce qui correspond à $t_y = d.t_z = 0$ (d est un coefficient de correction de perpendicularité qui est "presque" nul).

Puis : $l_2 = C_G \wedge C_D$ et $l_3 = (C_G - C_D) \wedge l_2 = (C_G - C_D) \wedge (C_G \wedge C_D)$. Ainsi, il nous reste 2 équations à 5 inconnues :

$$\begin{cases} l_1 \cdot C_G + l_{14} = 0 \\ l_1 \cdot C_D + l_{14} = \Delta_x \end{cases}$$

Pour éviter une matrice perspective dégénérée, nous prenons : $l_1 = (C_G \wedge C_D) \wedge (C_G + C_D)$.

En définitive :

$$M_G = \begin{pmatrix} & l_{14} & \\ Q_{33} & 0 & \\ & \| C_G \wedge C_D \|^2 & \end{pmatrix} \quad (6)$$

$$\text{et } M_D = \begin{pmatrix} & l_{14} - \Delta_x & \\ Q_{33} & 0 & \\ & \| C_G \wedge C_D \|^2 & \end{pmatrix}$$

$$\text{avec } Q_{33} = \begin{pmatrix} [(C_G \wedge C_D) \wedge (C_G + C_D)]^t \\ [(C_G \wedge C_D)]^t \\ [(C_G - C_D) \wedge (C_G \wedge C_D)]^t \end{pmatrix}$$

La rectification des images gauche et droite (en niveaux de gris) se fait à partir de la méthode décrite dans [10] combinée avec un algorithme de calcul de barycentres en fonction des 8 voisins pour ramener les coordonnées des points (en pixels) en valeurs entières.

4.2 Calcul des paramètres à partir de M_G et M_D

D'après (4) nous obtenons le système suivant :

$$\begin{cases} l_1 = \frac{1}{b}r_1 - \frac{c}{be}r_2 + \frac{(cd - ae)}{be}r_3 \\ l_2 = \frac{1}{e}r_2 - \frac{d}{e}r_3 \\ l_3 = r_3 \\ l_{14} = \frac{1}{b}t_x - \frac{c}{be}t_y + \frac{(cd - ae)}{be}t_z \\ l_{24} = \frac{1}{e}t_y - \frac{d}{e}t_z \\ l_{34} = t_z \end{cases} \quad (7)$$

A partir de ce système d'équations, nous pouvons appliquer directement la méthode de résolution décrite dans [10] pour le calcul des paramètres liés au modèle de formation d'image (a, b, c, d, e), pour la matrice de rotation \mathcal{R} et pour la translation T. En même temps nous déterminons la valeur du paramètre Δ_x de la translation T_{GD} .

5 Contrainte de planéité

L'utilisation de la contrainte de planéité a été faite suivant des approches différentes. D'une part, Takahashi [11] l'utilise comme critère local pour regrouper simultanément des segments 2-D. D'autre part, Vezien [12] l'utilise pour définir une transformation approchée du plan image gauche au plan image droit et réciproquement.

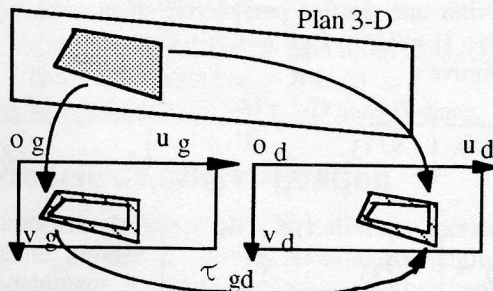


Figure 2 : Contrainte de planéité

Cette contrainte assure l'existence d'une transformation homographique entre les coordonnées de points homologues (u_g, v_g) et (u_d, v_d) provenant d'un même point 3-D situé dans un plan donné. Dans notre cas (géométrie rectifiée), nous obtenons une transformation homographique τ_{gd} qui se traduit par :

$$\begin{pmatrix} u_d & v_d \end{pmatrix}^t = \tau_{gd} \cdot \begin{pmatrix} u_g & v_g & 1 \end{pmatrix}^t ; \quad (8)$$

avec $\tau_{gd} = \begin{pmatrix} \alpha_u & \alpha_v & \lambda \\ 0 & 1 & 0 \end{pmatrix}$

Les coefficients α_u , α_v et λ de la transformation τ_{gd} sont fonction des paramètres du banc stéréoscopique et des paramètres définissant le plan 3-D (exprimés dans un repère donné).

5.1 Expression de τ_{gd} dans (G)

Tout plan parallèle à l'axe (O_G, Z_G) ne sera pas visible par la caméra, donc nous nous intéressons uniquement aux plans ne contenant pas cet axe. Ainsi,

l'équation du plan contenant la facette 3-D à reconstruire peut s'écrire sous la forme :

$$Z_G = p \cdot X_G + q \cdot Y_G + r \quad (9)$$

ou $1 = p \frac{X_G}{Z_G} + q \frac{Y_G}{Z_G} + \frac{r}{Z_G}$ (pour $Z_G \neq 0$)

Or, d'après (1) et (3) nous avons :

$$\begin{cases} \frac{X_G}{Z_G} = b u_g + c v_g + a \\ \frac{Y_G}{Z_G} = e v_g + d \end{cases} \quad (10)$$

et en combinant (9) et (10), nous déduisons :

$$\frac{1}{Z_G} = \frac{b}{\Delta_x} (u_g - u_d) \quad (11)$$

puis, en reportant (11) et (10) dans (9), nous obtenons :

$$u_d = \left[\frac{\Delta_x}{r} p + 1 \right] u_g + \frac{\Delta_x}{rb} [cp + eq] v_g + \frac{\Delta_x}{rb} [ap + dq - 1]$$

En définitive, nous avons :

$$\tau_{gd} = \begin{pmatrix} \left(\frac{\Delta_x}{r} p + 1 \right) & \frac{\Delta_x}{rb} (cp + eq) & \frac{\Delta_x}{rb} (ap + dq - 1) \\ 0 & 1 & 0 \end{pmatrix} \quad (12)$$

5.2 Disparité et gradient de disparité

a) Soient $P_g^1 = (u_g^1, v_g^1)$ un point de R_g et $P_d^1 = (u_d^1, v_d^1)$ son homologue dans R_d . La disparité d_u (horizontale seulement dans notre cas) s'exprime d'après (8) sous la forme d'une expression affine :

$$d_u = u_d^1 - u_g^1 = (\alpha_u - 1) u_g^1 + \alpha_v v_g^1 + \lambda \quad (13)$$

b) Le gradient de disparité δd_u , exprimé en fonction des coordonnées images des deux couples de points homologues, est :

$$\delta d_u = (\alpha_u - 1 \ \alpha_v) \cdot \begin{pmatrix} \Delta u_g \\ \Delta v_g \end{pmatrix}, \quad \text{où :} \quad (13\text{bis})$$

Δu_g : différence des abscisses ;
 Δv_g : différence des ordonnées

c) Nous en déduisons l'interprétation suivante :

- $\alpha_u - 1$: coefficient de linéarité de la disparité d_u dans la direction horizontale (axe des abscisses : u).
- α_v : coefficient de linéarité de la disparité d_u dans la direction verticale (axe des ordonnées : v).
- λ : coefficient de translation dans la direction horizontale.

5.3 Calcul de l'équation du plan 3-D

Supposons que l'équation du plan (\mathcal{P}) recherché s'écrive dans (G) sous la forme donnée dans (9). On a :

$$p = (\alpha_u - 1) \cdot \frac{r}{\Delta} \quad q = [b\alpha_v + c(1 - \alpha_u)] \cdot \frac{r}{\Delta e}$$

$$\text{et } \frac{1}{r} = \frac{1}{\Delta} \left[\frac{1}{e} (cd - ae) (1 - \alpha_u) + \frac{bd}{e} \alpha_v - b\lambda \right]$$

D'où, en utilisant (2), l'équation du plan (\mathcal{P}) dans le repère (M) est donnée par : $A X + B Y + C Z + D = 0$;

avec : $(A B C D)^t = \begin{pmatrix} \mathcal{R}^t & O_{31} \\ T^t & 1 \end{pmatrix} \begin{pmatrix} -p \\ -q \\ 1 \\ -r \end{pmatrix}$ (14)

où O_{ij} est la matrice nulle de dimension $i \times j$.

5.4 Implantation

Soit (R_g, R_d) un couple de régions homologues provenant de la même facette 3-D.

5.4.1 Définitions

a) *Définition 1*

Nous appellerons "segment épipolaire" associé à la région R_g (resp. R_d), l'intersection d'une droite épipolaire gauche (resp. droite) avec R_g (resp. R_d).

b) *Définition 2*

Nous appellerons "couple de segments épipolaires conjugués" associé à (R_g, R_d) , un couple formé de deux segments épipolaires associés respectivement à R_g et R_d et issus d'un couple de droites épipolaires conjuguées.

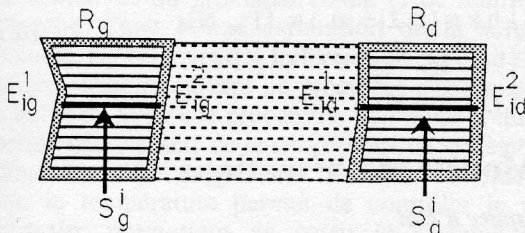


Figure 4 : Ensemble des segments épipolaires conjugués

c) *Notations*

- un couple de segments épipolaires conjugués sera noté par la suite par (S^i_g, S^i_d) avec $S^i_g = (E^1_{ig}, E^2_{ig})$ et $S^i_d = (E^1_{id}, E^2_{id})$ où les points 2-D, E^1_{ig} , E^2_{ig} , E^1_{id} et E^2_{id} sont les quatre extrémités des deux segments S^i_g et S^i_d .

- au couple de région (R_g, R_d) , nous associons les N couples des segments épipolaires conjugués (S^i_g, S^i_d) qui lui sont associés et l'ensemble des extrémités:

$$\mathcal{S}(R_g, R_d) = \{ (E^1_{ig}, E^1_{id}, E^2_{ig}, E^2_{id}) / i \in \{1 \dots N\} \}$$

- pour une extrémité E^h_{id} donnée, la zone d'incertitude due à la segmentation correspondante est notée : \mathcal{J}_{hi} .

5.4.2 Estimation de τ_{gd}

Soit $i \in \{1, \dots, N\}$ et (S^i_g, S^i_d) le couple de segments épipolaires conjugués numéro i .

a) Le gradient de disparité $(\delta d_u)_{ii}$ entre les deux extrémités est : $(\delta d_u)_{ii} = (\alpha_u - 1) \cdot (u^1_g - u^2_g)$. Ceci nous permet de calculer une estimation α_u^i du coefficient α_u à partir du couple (S^i_g, S^i_d) . Soit :

$$\alpha_u^i = \frac{(\delta d_u)_{ii}}{u^1_g - u^2_g} + 1 \quad (15)$$

b) Considérons aussi les deux couples de segments voisins (S^{i-1}_g, S^{i-1}_d) , (S^{i+1}_g, S^{i+1}_d) et un quadruplet des extrémités de la forme $(E^h_{ig}, E^h_{id}, E^k_{jg}, E^k_{jd})$, où $h, k \in \{1, 2\}^2$ et $j \in \{i-1, i+1\}$. Notons qu'il y a 8 quadruplets distincts pour un indice i donné. Par application de l'équation (13 bis), nous obtenons :

$$(\delta d_u)_{i,j} = (\alpha_u^i - 1) \cdot (u^h_{ig} - u^k_{jg}) + \alpha_v \cdot (v^h_{ig} - v^k_{jg}).$$

D'où nous déduisons une estimation $\alpha_v^i(h, k, j)$ de α_v , puis nous calculons la moyenne arithmétique des 8 estimations possibles comme meilleur estimateur de α_v^i , soit :

$$\alpha_v^i = \frac{1}{8} \sum_{h,k \in \{1,2\}} \sum_{j \in \{i-1, i+1\}} \alpha_v^i(h, k, j) \quad (16)$$

c) Enfin, par application de l'équation (13) pour l'une des deux extrémités du couple des segments épipolaires conjugués numéro i , nous déduisons une estimation λ^i du coefficient λ .

$$\lambda^i = u_d - \alpha_u^i \cdot u_g - \alpha_v^i \cdot v_g \quad (17)$$

5.4.3 Conclusion

Ainsi, à chaque couple de segments (S^i_g, S^i_d) nous avons associé une estimation τ^i_{gd} de la transformation $\tau_{gd}(R_g, R_d)$. Dans le cas idéal, il faudrait avoir :

$$\tau^i_{gd} = \tau^j_{gd} \cdot \forall (i, j) \in \{1, \dots, N\}^2 \quad (18)$$

Nous utiliserons cette condition pour corriger la frontière de l'une des deux régions (R_g et R_d) en fonction de l'autre. En d'autres termes, sachant que la segmentation en régions est instable, nous essayons de prédire une frontière fictive localisée autour de la région droite R_d qui correspondrait au mieux à celle de la région gauche R_g . En effet, nous chercherons à associer une transformation unique à tous les couples de segments épipolaires conjugués (S^i_g, S^i_d) (où $i \in \{1, \dots, N\}$). Ceci est possible par la minimisation de la fonction suivante :

$$E_1(R_g, R_d) = 3 - [\rho(\alpha_u) + \rho(\alpha_v) + \rho(\lambda)] \quad (19)$$

$$\text{avec : } \begin{cases} \rho(\alpha_u) = \frac{\text{Min}(\alpha_u^i)}{\text{Max}(\alpha_u^i)} \\ \rho(\alpha_v) = \frac{\text{Min}(\alpha_v^i)}{\text{Max}(\alpha_v^i)} \\ \rho(\lambda) = \frac{\text{Min}(\lambda^i)}{\text{Max}(\lambda^i)} \end{cases} \quad \text{où } i \in \{1 \dots N\}$$

6 Contrainte radiométrique

Nous utiliserons aussi des informations radiométriques pour mesurer la similarité entre les pixels de l'image gauche avec ceux de l'image droite. Vu que la comparaison des niveaux de gris de deux pixels est assez sensible au bruit dans les images, nous utiliserons la technique de corrélation. Ainsi, nous

pouvons comparer les pixels gauches et droites en considérant une fenêtre centrée sur le pixel gauche et une centrée sur le pixel droit. Nous choisissons la meilleure mise en correspondance au niveau de la similarité radiométrique grâce à cette corrélation. En effet, le coefficient de corrélation est maximal si la similarité est maximale entre les deux fenêtres gauche et droite. Notons aussi, que ce critère est plus efficace quand le niveau de gris est non uniforme dans la zone considérée de l'image.

6.1 Définitions

Soit $p(i,j)$ un pixel donné.

- $V_k(i,j) = \{ (i',j') / |i - i'| \leq k \text{ et } |j - j'| \leq k \}$: la fenêtre de taille $[2k + 1, 2k + 1]$ définissant le voisinage du pixel $p(i,j)$ à l'ordre k . Nous prenons par la suite $k = 1$.
- $X_g(i',j')$: la valeur du niveau de gris du pixel $p(i',j')$ de l'image gauche.
- $X_d(i',j')$: la valeur du niveau de gris du pixel $p(i',j')$ de l'image droite.
- Nous considérons X_l et X_r comme deux variables aléatoires uniformes définies sur le domaine $V_k(i,j)$. D'où, pour tout couple (i',j') appartenant à $V_k(i,j)$ la probabilité pour que $X_l(i',j') = X_g(i',j')$ est : $\frac{1}{(2k + 1)^2}$

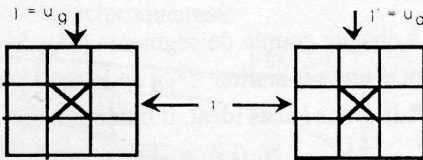


Figure 4: Coefficient de corrélation entre deux pixels homologues

6.2 Similarité entre pixels homologues

Soient (u_g, v_g) et (u_d, v_d) deux pixels homologues appartenant au couple des segments épipolaires conjugués numéro i . Par conséquent, $v_g = v_d = Cte_i$. Le coefficient de corrélation $r(u_g, u_d, i)$ est défini par :

$$r(u_g, u_d, i) = \frac{\mu_{11}}{\sigma_g \sigma_d}; \text{ avec :}$$

$$\begin{cases} \mu_{11} : \text{Covariance de } X_g \text{ et } X_d \\ \sigma_g : \text{Ecart type de } X_g \\ \sigma_d : \text{Ecart type de } X_d \end{cases}$$

On a la propriété : $\forall (i, u_g, u_d) : |r(u_g, u_d, i)| \leq 1$. Donc, pour mesurer la similarité entre les pixels gauche et droit (u_g, v_g) et (u_d, v_d) , il suffit de minimiser la fonction :

$$\gamma(u_g, u_d, i) = 1 - |r(u_g, u_d, i)| \quad (20)$$

6.3 Similarité entre segments homologues

Etant donné un couple de segments épipolaires (S^i_g, S^i_d) , nous lui associons la fonction coût suivante :

$$\gamma^i = \sum_{\substack{(u_g, v_g) \in [E^1_{ig}, E^2_{ig}] \\ (u_d, v_d) \in [E^1_{id}, E^2_{id}]}} \gamma(u_g, u_d, i) \quad (21)$$

6.4 Similarité entre régions homologues

Puis, nous associons la fonction énergie $E_2(R_g, R_d)$ à chaque couple de régions homologues définie de la façon suivante :

$$E_2(R_g, R_d) = \sum_{i \in \{1 \dots N\}} \gamma^i \quad (22)$$

6.5 Fonction énergie

Pour chaque couple de régions homologues (R_g, R_d) , nous considérons les N couples de segments épipolaires conjugués qui lui sont associés et l'ensemble $\mathcal{P}(R_g, R_d)$. D'où, nous définissons la fonction énergie suivante à partir de la contrainte de planéité et de la corrélation radiométrique :

$$\begin{cases} E = \lambda_1 E_1(R_g, R_d) + \lambda_2 E_2(R_g, R_d) \\ h, k \in \{1, 2\}^2 \text{ et } i \in \{1 \dots N\} \\ (E^h_{ig}, E^k_{id}) \in \mathcal{P}(R_g, R_d) \\ (\lambda_1, \lambda_2) \in (\mathbb{R}^+)^2 \end{cases} \quad (23)$$

7 Modèle mathématique

a) Espace d'état

L'espace des états est :

$S = \{ s(h, i, u_d) = P^h_{i}(u_d) / h \in \{1, 2\}, i \in \{1 \dots N\}, u_d \in \mathcal{J}_{hi} \}$. S représente l'état des correspondances des extrémités gauche-droite de la forme (E^h_{ig}, E^h_{id}) et la localisation de l'extrémité E^h_{id} dans l'intervalle \mathcal{J}_{hi} correspondant. Où :

$$P^h_{i}(u_d) = \begin{cases} 1 & \text{si } E^h_{ig} \text{ est apparié à } E^h_{id} \\ & \text{et l'abscisse de } E^h_{id} \text{ est } u_d \\ 0 & \text{si non} \end{cases}$$

b) Contrainte d'unicité

La contrainte d'unicité pour la mise en correspondance des extrémités est assurée par la propriété (U) suivante :

$$(U) : \begin{cases} \sum P^h_{i}(u_d) = 1 \\ u_d \in \mathcal{J}_{hi} \\ \forall i \in \{1 \dots N\} \text{ et } \forall h \in \{1, 2\} \end{cases}$$

c) Processus stochastique

Nous associons à S un processus stochastique $\{P_s\}_{s \in S}$. $S = s(h, i, u_d)$ est un état réalisant la variable aléatoire P_s (à valeur dans $\{0, 1\}$), pour le couple d'extrémités homologues (E^h_{ig}, E^h_{id}) et à \mathcal{J}_{hi} .

d) Problème de minimisation avec contrainte

En intégrant les différentes informations obtenues précédemment, le problème de recalage des frontières

peut se formaliser sous forme du problème de minimisation (P) suivant :

$$(P) : \begin{cases} \text{Min } E(P^{h_i}(u_d)) = \sum_{j=1}^2 \lambda_j E_j(R_g, R_d) \\ h \in \{1, 2\} \text{ et } i \in \{1 \dots N\} \\ P^{h_i}(u_d) \in (U) \text{ et } u_d \in \mathcal{J}_{hi} \\ \lambda_j \in \mathbb{R}^+ \end{cases}$$

8 Résolution et recuit simulé

8.1 Le recuit simulé

Cet algorithme se fonde sur une analogie entre les problèmes d'optimisation et ceux de la physique statistique [6]. En effet, il permet de considérer un petit échantillon parmi toutes les configurations microscopiques possibles d'un système et de calculer des propriétés macroscopiques en fonction de la température.

Ainsi, en optimisation combinatoire, cette méthode nous permet de n'explorer que certaines configurations représentatives du problème étudié et de réaliser la recherche d'une bonne estimation de la solution minimale. Le parcours des différents états se fait d'une manière aléatoire suivant un processus d'optimisation qui autorise, avec une probabilité non nulle et qui est fonction décroissante de la température du système, un changement d'état augmentant la fonction énergie. En effet, la température permet de contrôler le coût nécessaire permettant de sortir le système d'un minimum local. Notons, qu'avec une température nulle l'algorithme opère comme toute technique strictement décroissante.

8.2 Implantation

Pour l'implantation de cet algorithme, il est nécessaire de préciser les 5 étapes suivantes :

a) Système d'état et état initial

Dans notre cas il suffit de considérer un couple de régions homologues (R_g, R_d) et de remplacer le système physique par l'ensemble $\mathcal{P}(R_g, R_d) = \{ (E^1_{ig}, E^1_{id}, E^2_{ig}, E^2_{id}) / i \in \{1 \dots N\} \}$ des extrémités de couples de segments épipolaires conjugués associés à (R_g, R_d). Ainsi, l'énergie du système pour un état donné est calculée par l'équation (17). Nous fixons comme état initial la configuration correspondant aux frontières déterminées par la segmentation (l'ensemble $\mathcal{P}(R_g, R_d)$).

b) Décroissance de la température

Nous fixons une température initiale t_0 assez élevée pour permettre à l'algorithme de sortir des minima locaux. Nous appliquons l'algorithme de Metropolis [6] qui produit une chaîne de Markov (suite d'états à température constante). Puis, nous faisons décroître la température pour une nouvelle exploration des états du système. La décroissance de la température t prend successivement les termes de la suite définie par :

$$u_0 = t_0 \text{ et } u_n = \frac{t_0}{1+n}$$

c) Changement d'état

Un changement d'état consiste à tirer au hasard une extrémité droite $E^{h_{id}}$ et à la déplacer aléatoirement à l'intérieur de la zone d'incertitude \mathcal{J}_{hi} (définie par $\mathcal{J}_{hi} = [u_d - \Delta n, u_d + \Delta n]$). Ceci revient à choisir au hasard trois entiers : $i \in \{1 \dots N\}$, $h \in \{1, 2\}$ et $u_d' \in \mathcal{J}_{hi}$. Puis il suffit de faire le changement suivant :

$$\begin{cases} P^{h_i}(u_d) = 1 - P^{h_i}(u_d) \\ P^{h_i}(u_d') = 1 - P^{h_i}(u_d') \end{cases}$$

d) Equilibre thermique

Nous supposons que l'équilibre thermique à une température donnée soit atteint dès qu'un changement d'état est refusé par l'algorithme. Nous constatons que la longueur de la chaîne de Markov obtenue à une température donnée est en moyenne égale à 1.

e) Crière d'arrêt

L'arrêt de l'algorithme se produit dès que nous atteignons une température "presque nulle" ou dès que l'équilibre du système total est atteint (300 itérations sans changement d'état).

9 Expérimentation

Les résultats obtenus pour les images de scènes d'intérieur prises par le robot de l'INRIA, nous permettent de dire que l'algorithme améliore la localisation des frontières des régions qui vérifient l'hypothèse faite sur l'incertitude de détection. Nous arrivons à reconstruire les sources lumineuses situées au plafond et la table (plans horizontaux) avec une erreur très faible. Mais, nous n'avons pas obtenu de résultats satisfaisants pour la reconstruction des petites régions. Pour les plans obliques (carreaux de fenêtres), nous avons actuellement une erreur inférieure à 10 degrés sur la détermination de la normale du plan. Notons aussi que cet algorithme permet de mieux reconstruire des facettes partiellement cachées (exemple: la table).

Le temps d'exécution est inférieur à une minute pour une température t_0 inférieure à 50°.

Type de la surface plane	Erreur maximale sur l'orientation de la normale
horizontale	2 degrés
oblique	10 degrés

Sur la page suivante se trouvent la mise en correspondance d'un couple d'images segmentées en régions (figure 5) et le résultat de la procédure de rectification sur le même couple d'images (figure 6). Enfin, la figure 7 montre un exemple de recalage des frontières de régions pour quelques couples de régions homologues.

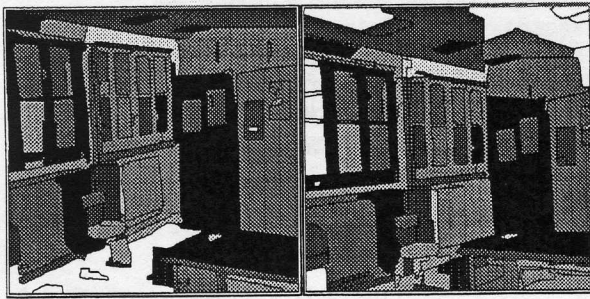


Figure 5 : Mise en correspondance des régions

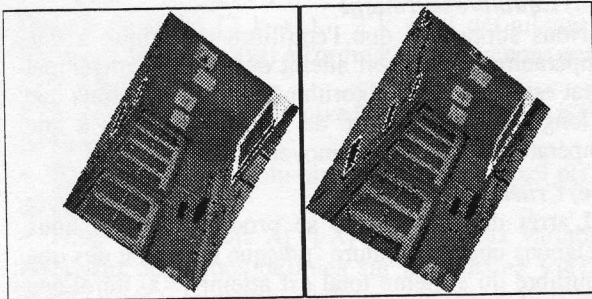


Figure 6 : Paire d'images de niveaux de gris rectifiées

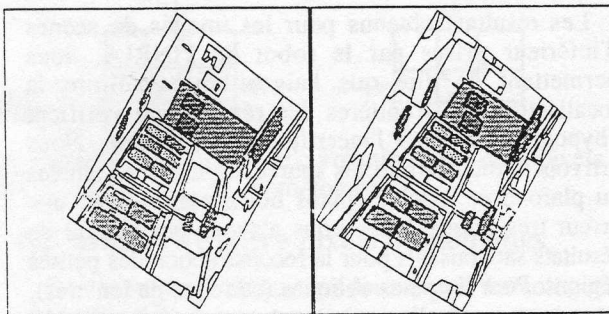


Figure 7 : Recalage des frontières de régions

10 Conclusion

Nous avons présenté une technique de reconstruction 3-D par analyse des primitives 2-D "région" et qui tient compte de la limitation majeure d'une telle primitive en essayant de réaliser une correction de la carte de disparité issue des deux phases précédentes (segmentation et appariement). Nous avons modéliser la rectification d'images afin de faciliter l'utilisation de la contrainte de planéité. Cette contrainte, combinée avec une contrainte radiométrique (corrélation) nous a permis de définir une fonction énergie dont la minimisation nous fournit le plan 3-D support de la facette 3-D recherchée. Nous déterminons les frontières de la facette 3-D dans l'espace en calculant l'intersection du plan trouvé avec les différents rayons de projection partant de la frontière de l'une des deux régions homologues.

Les résultats obtenus sont très satisfaisants pour les régions correctement appariées et dont la segmentation est satisfaisante. De plus, cette méthode permet la

reconstruction de facettes partiellement cachées exceptées les régions situées au bord de l'image.

Références

- [1] A. Abi-Ayad - B. Thiesse and All. "Les multiples facettes du problème de calibrage de caméras". In Proc. 7ème congrès R.F.I.A., pp. 507 - 525, Paris, France, November 1989.
- [2] N. Ayache : "Vision stéréoscopique et perception multisensorielle : Application à la robotique mobile". Collection Science Informatique, InterEditions, France, 1989.
- [3] P. Bonnin : "Méthode systématique de conception et de réalisation d'applications en vision par ordinateur". PhD thesis, Université de Paris VII, 1991.
- [4] S.Y. Chen and al. : "Split-and-Merge Image Segmentation Based on Localized Feature Analysis and Statistical Tests". CVGIP-53-5, pp 457-475, September 1991.
- [5] W.E.L. Grimson : "Computational Experiments with a Feature Based Stereo Algorithm", IEEE PAMI-7-1, pp 17-34, Janvier 1985.
- [6] D.S. Johnson, C.R. Aragon, L.A. McGeoch, and C. Schvon. "Optimization by simulated annealing : an experimental evaluation ; part ii, graph coloring and number partitioning". Operations Research, 39(3) : 378-406, 1991.
- [7] R. Mohan et R. Nevatia : "Using Perceptual Organization to Extract 3-D Structures". IEEE Trans. Pattern Anal. Machine Intell. vol. 11, n°11, 1989, pp1121-1139.
- [8] O. Monga and B. Wrobel "Segmentation d'images : vers une méthodologie". Traitement du signal, 4(3) : 169, 1987. Special issue on Computer Vision.
- [9] T. Skordas " Mise en correspondance et reconstruction stéréo, utilisant une description structurelle des images ". PhD thesis, I.N.P.G., France, 1988.
- [10] G. Toscani "Systèmes de calibration et perception du mouvement en vision artificielle". PhD thesis, Université de Paris-Sud, 1987.
- [11] H. Takahashi, F. Tominta : "Planarity constraint in stereo Matching". 9th int Conf. on Pattern Recognition Rome, November 1988, pp. 446-449.
- [12] Vezien J.M. - Gagalowitz A. " Reconstruction 3-D basée sur une analyse en régions d'une paire d'images stéréoscopiques". 8ème congrès R.F.I.A., pp. 649 - 658, France, November 1991.
- [13] B. Wrobel - Dautcourt. "Perception de la distance par mise en correspondance de régions entre des images stéréoscopiques". PhD thesis, I.N.P.L., France, March 1988.
- [14] E. Zagrouba and C. Krey : "Region matching by adjacency propagation in stereovision". ICARCV'92, International Conference on Automation, Robotics and Computer. Vision, Volume 1, pp. CV-8.5.1.CV-8.5.5. 16-18 September 1992, SINGAPORE.
- [15] E. Zagrouba, M. Buvry and C. Krey : "A Rule-Based System for Region Segmentation Improvement in Stereovision". IS&T/SPIE Symposium on Electronic Imaging, Science & Technology, 6-10 February 1994, San Jose, California, USA.