



# Reconstruction globale et robuste de facettes 3D

Jean-Philippe Tarel

► **To cite this version:**

Jean-Philippe Tarel. Reconstruction globale et robuste de facettes 3D. [Rapport de recherche] RR-2813, INRIA. 1996, pp.38. <inria-00073879>

**HAL Id: inria-00073879**

**<https://hal.inria.fr/inria-00073879>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Reconstruction globale et robuste de facettes 3D***

Jean-Philippe Tarel

**N° 2813**

Février 1996

PROGRAMME 4

  
*Rapport  
de recherche*



## Reconstruction globale et robuste de facettes 3D

Jean-Philippe Tarel \*

Programme 4 — Robotique, image et vision  
Projet Syntim

Rapport de recherche n° 2813 — Février 1996 — 38 pages

**Résumé :** Nous présentons une approche de la reconstruction 3D peu explorée jusqu'à présent qui présente pourtant des avantages importants. Dite globale, elle s'attache à reconstruire des primitives géométriques de hauts niveaux en s'appuyant uniquement sur les caractéristiques géométriques globales des projections 'image' de la primitive. Notre primitive de base est la facette 3D, et les moments géométriques des régions images sont nos caractéristiques. Notre méthode présente les avantages suivants :

- l'utilisation des moments géométriques des régions images ne nécessite aucune mise en correspondance au niveau du pixel et rend la méthode robuste aux petites erreurs de segmentation sur leurs contours,
- les paires de régions peuvent être sélectionnées pour ne reconstruire que les primitives 3D qui semblent valides, ce qui permet de les séparer des faces probablement occultées,
- les faces occultées sont traitées par une méthode globale fondée sur la photométrie des régions.

Enfin, un outil de validation de méthodes de reconstruction est présenté et utilisé sur notre approche.

**Mots-clé :** géométrie 3D, reconstruction, facette planes, disparités, moments, rectification, occlusions.

*(Abstract: pto)*

\*. E-mail : Jean-Philippe.Tarel@inria.fr

# Global and Robust Reconstruction of 3D Planar Patches

**Abstract:** We present a 3D reconstruction approach which is few studied in consideration of its numerous advantages. It is a global approach, where the algorithm attempt to reconstruct geometric high level features by using only global parameters of the image projections of the feature. Our elementary feature is 3D planar patch, and geometric moments are our global parameters. The proposed method of reconstruction has the following adavantages:

- using of the geometric moments of image region do not need a pixel-to-pixel matching and yields robustness to small errors of segmentation on region edges,
- valid matched regions are selected and thus, probably occluded planar patches are eliminated,
- occluded paches are processed by a global reconstruction algorithm based on region photometry.

Finally, a validation tool of reconstruction method is presented and used on our results.

**Key-words:** 3D geometry, reconstruction, planar patches, disparities, moments, rectified cameras, occlusions.

# 1 Introduction

La reconstruction permet d'obtenir des informations 3D sur la scène observée, à partir de données extraites dans les images. De fait, elle nécessite des étapes d'analyse préliminaires, dont le nombre et la complexité dépendent du type de primitives reconstruites. Généralement, trois types de processus sont effectués sur la paire stéréoscopique :

- La calibration de la paire corrige d'une part les distorsions de la caméra pour se placer dans le modèle sténopé parfait, et d'autre part calcule le déplacement entre les deux points de vue [TG95],
- Les primitives contenues dans chaque image sont extraites de celles-ci. Pour notre système de vision, les informations extraites sont des régions. L'algorithme d'extraction utilisé est alors une méthode de segmentation, où l'image est partitionnée en régions supposées être la projection d'une même facette d'un objet de l'espace 3D [Vin91, Ran92, AMG93, BSG94].
- Enfin, les primitives extraites dans l'image gauche et droite sont mises en correspondance deux à deux [SVCG89, Ran92].

La primitive géométrique la plus simple pour faire de la photométrie est la facette 3D. Comme nous désirons poursuivre la reconstruction géométrique par une reconstruction photométrique, nous avons choisi la région comme primitive de base dans l'image. Il existe de nombreux systèmes de vision qui effectuent un autre choix de primitive tels les points [MF92, BMV93], les segments [Aya88], et les courbes [RZMF92, Ma93, Ber95]. Mais, les primitives de plus bas niveau, comme les points ou les segments, bien que plus faciles à extraire, nécessitent trop souvent une étape supplémentaire, complexe, pour les réunir sous forme de facettes en 3D et d'objets.

Or, le choix des régions comme primitives extraites des images pose le problème de reconstruire des facettes 3D à partir d'une paire de régions (figure 1). Chacune de ces dernières est la projection d'une face plane de l'espace 3D. Le problème est donc, avec des appariements de régions de l'image gauche et droite, de retrouver la position des supports des différentes faces 3D présentes dans la scène.

Citons les méthodes de reconstruction de facettes 3D. D'abord, il est possible de travailler localement en utilisant explicitement des primitives de plus bas niveau que

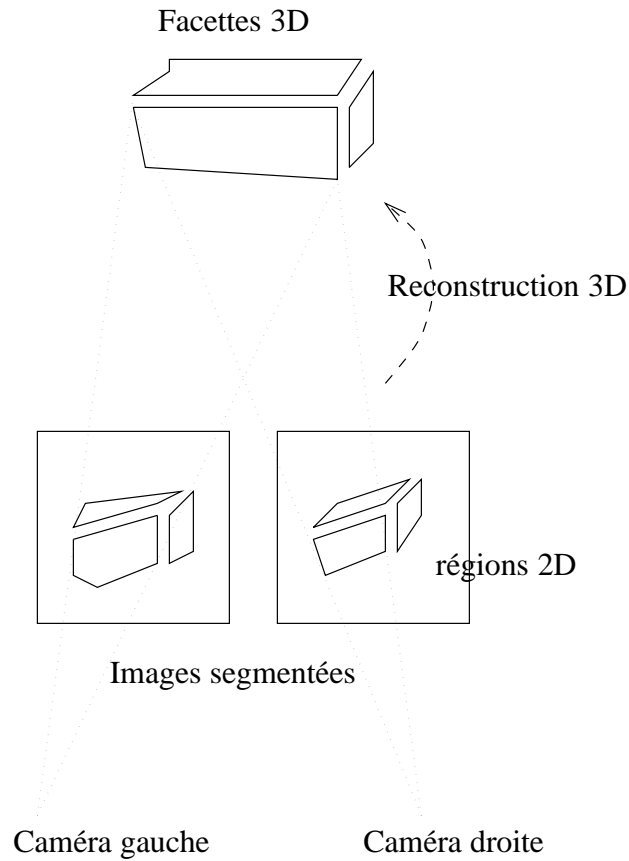


FIG. 1 - *Reconstruction d'un objet 3D à partir d'une paire segmentée en régions.*

les régions, tels les sommets qui caractérisent ses contours. Par exemple, une fois mis en correspondance, grâce à la contrainte épipolaire, les sommets peuvent être reconstruits par triangulation. Comme l'information sur l'appartenance des points à la facette est conservée, la facette est déduite. Il est possible d'utiliser de façon analogue les segments qui forment l'approximation polygonale du contour [CM94, JF92].

Une deuxième approche, plus originale, utilise les informations globales sur la paire de régions. Si la transformation perspective réalisée par la caméra est ap-

proximée par une transformation orthographique ou para-perspective [Alo90] qui est plus simple, la position de la facette s'exprime en fonction des caractéristiques géométriques des régions gauche et droite [CO85, Véz95, VT95, TV95]. Un raffinement itératif par minimisation des erreurs introduites dans l'approximation est aussi réalisable [LVG92]. Néanmoins, l'aspect itératif de cette méthode conduit souvent à des temps de calculs longs, avec le risque d'être piégé dans un minimum local.

L'avantage des méthodes locales sur celles globales est de traiter les occultations de la facette. En revanche, les méthodes locales sont généralement peu robustes aux inévitables erreurs de segmentation.

La méthode exposée dans ce rapport, qui évite toute approximation de la transformation perspective, s'inscrit dans le cadre d'une approche globale précise et robuste aux petites erreurs de segmentation. Bien que le problème soit non linéaire, nous avons cherché une méthode directe de reconstruction des facettes 3D. Ainsi, nous proposons de rectifier les images. En effet, à cette seule condition, les équations, qui lient les caractéristiques géométriques des régions et le plan support de la face tridimensionnelle, sont utilisables. Cette étape préliminaire de rectification est réalisable juste après la calibration, et elle rend les épipolaires parallèles entre elles sur les deux images. Alors, en se restreignant aux images rectifiées, deux méthodes de reconstruction globale de faces planes sont proposées, l'une géométrique, l'autre photométrique.

Lors de cette étude, il nous a semblé utile de travailler dans un espace intermédiaire dénommé *espace des disparités*. En effet, les équations qui lient le 3D aux projections 2D y sont plus simples que celles écrites dans l'espace réel. Il est donc plus facile de raisonner dans l'*espace des disparités*. Le retour à l'espace 3D métrique ne peut alors être réalisé que si l'on dispose de la calibration de la paire. Si cette dernière n'est pas connue, comme dans certaines approches [MF92, BMV93], on peut exploiter cet espace non métrique qui est isomorphe au réel à une transformation homographique près.

Après avoir récapitulé les relations entre un point et ses projections suivant qu'il est libre ou contraint d'appartenir à un plan fixe, nous introduisons dans la section 2 l'espace intermédiaire de travail dit *espace des disparités*, qui permet de simplifier de beaucoup les équations. L'existence de cet espace nécessite une rectification de la géométrie des caméras. Dans la section 3, nous mettons à profit l'*espace des disparités* pour résoudre la reconstruction photométrique comme géométrique de facettes 3D. Cette dernière est robuste et précise mais nécessite des faces non



occultées. Or la présence d'une occultation peut être partiellement testée avec des critères globaux. Nous décrivons donc de quelle manière les faces non occultées sont sélectionnées. De façon complémentaire, la reconstruction photométrique est robuste aux occultations, mais nécessite la présence de dégradés réguliers d'éclairage sur les régions. Dans la section 4, les résultats produits par notre méthode sont présentés, qui illustrent sa robustesse et sa capacité à traiter les occultations. Enfin dans la section 5, la précision de la méthode est discutée et comparée avec d'autres approches.

## 2 L'espace des disparités

### 2.1 Projection stéréoscopique d'un point

En toute généralité, lors de l'acquisition d'une paire stéréoscopique, la position des caméras est quelconque. Le lien qui unit les coordonnées  $X$  d'un point de l'espace 3D, à son image gauche  $U_g = (u_g, v_g)$  et droite  $U_d = (u_d, v_d)$  est décrit à partir des relations suivantes :

$$\left. \begin{aligned} u_g &= \frac{L_{1g} \begin{pmatrix} X \\ 1 \end{pmatrix}}{L_{3g} \begin{pmatrix} X \\ 1 \end{pmatrix}} & u_d &= \frac{L_{1d} \begin{pmatrix} X \\ 1 \end{pmatrix}}{L_{3d} \begin{pmatrix} X \\ 1 \end{pmatrix}} \\ v_g &= \frac{L_{2g} \begin{pmatrix} X \\ 1 \end{pmatrix}}{L_{3g} \begin{pmatrix} X \\ 1 \end{pmatrix}} & v_d &= \frac{L_{2d} \begin{pmatrix} X \\ 1 \end{pmatrix}}{L_{3d} \begin{pmatrix} X \\ 1 \end{pmatrix}} \end{aligned} \right\} \quad (1)$$

La matrice de transformation perspective  $\mathcal{L}_i$  de la caméra  $i$  est constituée des lignes à quatre éléments  $L_{1i}$ ,  $L_{2i}$  et  $L_{3i}$ . En réunissant ces équations, on obtient la liaison entre les trois coordonnées  $X = (x, y, z)$  d'un point de l'espace et les quatre coordonnées  $(u_g, v_g, u_d, v_d)$  de son image dans la paire.

Un point est reconstruit par triangulation de sa position 3D à partir des coordonnées de ses deux projections 2D. La donnée de trois des coordonnées,  $(u_g, v_g, u_d)$  par exemple suffit puisque les coordonnées 'image' sont liées par la contrainte épipolaire (figure 2). Plus généralement, on peut établir un lien univoque entre les points 3D devant les deux rétines des caméras et les points de coordonnées  $(u_g, v_g, u_d)$

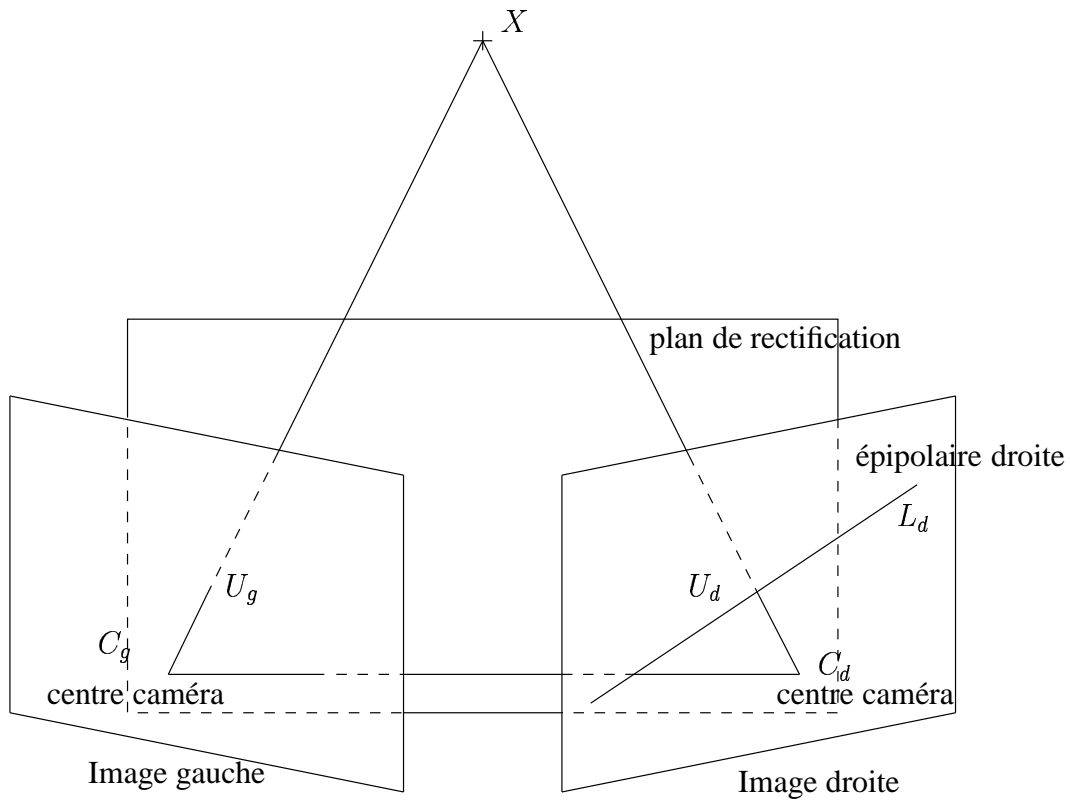


FIG. 2 - Le point  $U_g$  est l'image d'un point sur l'axe qui passe par  $C_g$  et  $X$ . Cette droite se projette sur l'image droite et définit la ligne épipolaire du point  $U_g$ . Le point  $U_d$  appartient donc nécessairement à la ligne épipolaire  $L_d(U_g)$ .

s'ils sont définis. Nous nommons *espace des disparités*, cet espace constitué des coordonnées  $(u_g, v_g, u_d)$  en bijection avec les points visibles par les caméras.

## 2.2 Projection stéréoscopique d'un point dans un plan

Dans le cas d'une géométrie des caméras quelconque, il est important d'étudier la transformation subie par une surface plane de l'espace réel, par passage dans

l'espace des disparités. Comme la relation entre l'espace réel et des disparités n'est pas linéaire, il est difficile de reconstruire des facettes 3D planes dans le cas général.

Soit un plan réel caractérisé par un point  $X_0$  et deux axes  $A_1$  et  $A_2$ . Avec les notations homogènes, un point  $X$  appartient au plan précédemment défini à la condition qu'il existe  $\lambda$  et  $\mu$  tels que :

$$\begin{pmatrix} X \\ 1 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & X_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ 1 \end{pmatrix} = \mathcal{P} \begin{pmatrix} \lambda \\ \mu \\ 1 \end{pmatrix}$$

La matrice  $3 \times 4$   $\mathcal{P}$ , ainsi introduite, caractérise le plan réel avec les notations homogènes. Ainsi, un point du plan, de coordonnées  $(\lambda, \mu)$ , a pour projection le point suivant, sur chaque image  $i$  :

$$U_i \equiv \mathcal{L}_i \mathcal{P} \begin{pmatrix} \lambda \\ \mu \\ 1 \end{pmatrix}$$

Le signe  $\equiv$  indique l'égalité à un facteur près. On retrouve l'énoncé comme quoi le lien entre les deux projections d'un point est une même homographie pour les points inclus dans un plan donné [Lus87] :

$$U_d \equiv (\mathcal{L}_d \mathcal{P})(\mathcal{L}_g \mathcal{P})^{-1} U_g \equiv \mathcal{H}_{d/g} U_g \quad (2)$$

La matrice  $3 \times 3$  de passage entre les points gauches et droits  $\mathcal{H}_{d/g}$  est homogène. À partir du premier terme de cette équation, nous déduisons qu'un plan 3D réel se transforme, par passage dans l'espace des disparités, en une quadrique d'équation :

$$u_d H_{3d/g} U_g - H_{1d/g} U_g = 0 \quad (3)$$

où  $H_{1d/g}$ ,  $H_{2d/g}$  et  $H_{3d/g}$  sont les lignes de la matrice  $\mathcal{H}_{d/g}$ . Les termes de plus haut degré de la quadrique précédente sont  $u_d u_g$  et  $u_d v_g$ .

La mise en œuvre d'une méthode de reconstruction 3D non fondée sur des appariements explicites nécessite la manipulation d'attributs globaux sur des ensembles de points 'image'. Les attributs géométriques les plus souvent utilisés sont les moments géométriques d'ordre petit, comme le barycentre et les moments d'inertie, puisqu'ils sont plus robustes au bruit que ceux d'ordre élevé. Toutefois, un plan de l'espace 3D est transformé en une surface complexe de l'espace des disparités qu'il

est difficile de manipuler. Par là même, dans le cas d'une géométrie des caméras quelconque, il n'est pas possible de relier explicitement les caractéristiques géométriques des régions gauche et droite en fonction des caractéristiques de la facette source.

Ainsi, pour reconstruire en 3D les faces de manière exacte grâce aux moments, il faut particulariser la géométrie des caméras gauche et droite. L'idée est donc d'imposer qu'un plan réel se transforme en un autre plan dans l'espace des disparités. Or, pour que l'équation de la quadrique (3) soit réduite à un plan, cela implique la condition suivante :

$$H_{3d/g} \equiv (0 \quad 0 \quad 1)$$

En utilisant l'équation (2), cette condition induit :

$$\mathcal{L}_d \mathcal{P} \equiv \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & 1 \end{pmatrix} \mathcal{L}_g \mathcal{P}$$

De cette équation, on déduit que  $L_{3g} = L_{3d}$ . Ceci traduit que les plans rétinien des deux caméras doivent être confondus. En conséquence, la géométrie des caméras doit être rectifiée pour permettre une reconstruction globale des faces, à la fois exacte et simple.

### 2.3 De l'espace des disparités vers celui réel

Nous allons à présent expliciter le lien entre un point réel et son correspondant dans l'espace des disparités.

Vu la contrainte  $L_{3g} = L_{3d} = L_3$ , le système d'équations (1) se simplifie et permet d'exprimer les coordonnées  $X$  d'un point réel en fonction des coordonnées 'image'  $(u_g, v_g, u_d)$  de manière linéaire :

$$\left. \begin{aligned} (u_g L_3 - L_{1g}) \begin{pmatrix} X \\ 1 \end{pmatrix} &= 0 \\ (v_g L_3 - L_{2g}) \begin{pmatrix} X \\ 1 \end{pmatrix} &= 0 \\ (u_d L_3 - L_{1d}) \begin{pmatrix} X \\ 1 \end{pmatrix} &= 0 \end{aligned} \right\} \quad (4)$$

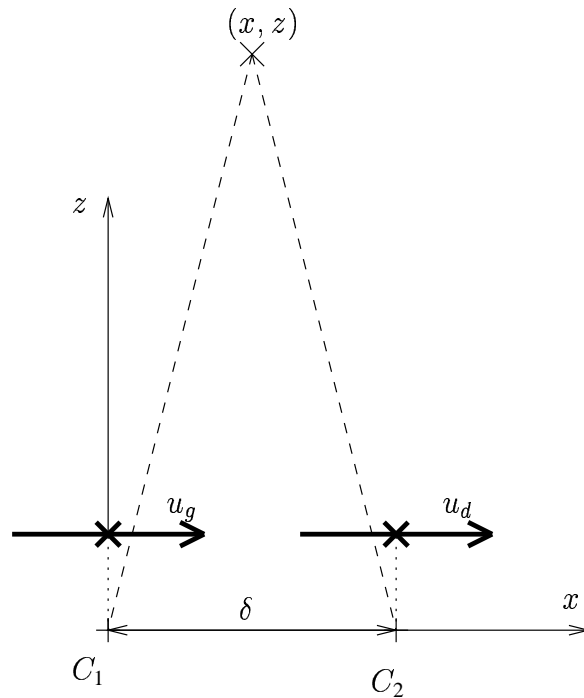


FIG. 3 - Géométrie des caméras rectifiées sur une coupe 2D.

Détaillons cette relation. Pour alléger les notations, nous choisissons l'origine des deux repères 'image' au centre des images (intersection de l'axe de visée de la caméra avec l'image), l'axe des  $x$  aligné avec les deux centres des caméras, et la longueur des axes égale à l'unité de l'espace 3D et non à la taille d'un pixel. Alors, sous ses hypothèses, on a  $v_g = v_d = v$  et les coordonnées verticale gauche  $u_g$ , verticale droite  $u_d$  et horizontale  $v$  d'un point de coordonnées  $X = (x, y, z)$  dans le repère de la caméra gauche sont :

$$\left. \begin{aligned} u_g &= \frac{x}{z} \\ u_d &= \frac{x - \delta}{z} \\ v &= \frac{y}{z} \end{aligned} \right\} \quad (5)$$

La distance entre les deux caméras, suivant l'axe des  $x$ , est donnée par  $\delta$  (voir figure 3). La transformation inverse est donc :

$$\left. \begin{aligned} x &= \frac{u_g \delta}{u_g - u_d} \\ y &= \frac{v \delta}{u_g - u_d} \\ z &= \frac{\delta}{u_g - u_d} \end{aligned} \right\} \quad (6)$$

L'inverse de la disparité  $u_d - u_g$  intervient directement dans cette transformation inverse, d'où la dénomination d'*espace des disparités*. La relation entre l'espace réel et des disparités est homographique d'après les équations (5) et (6). Cet espace est parfois introduit dans les méthodes de reconstruction par corrélation des pixels sans être véritablement dénommé : *search space for intra-scanline search* [OK85], *cross corrélation volume* [CM85]. L'introduction de l'espace des disparités facilite grandement la manipulation des facettes planes dans le cadre de la stéréovision. Effectivement, l'idée directrice est alors de travailler dans cet espace pour effectuer la reconstruction puis, grâce à la calibration du couple, retourner vers l'espace 3D réel. Travailler uniquement dans l'espace des disparités, sans faire le retour dans l'espace réel, se rapproche des méthodes de reconstruction utilisées dans le cadre de la recherche sur l'auto-calibration, puisque la reconstruction est alors obtenue à une transformée perspective près. De fait, si l'on dispose d'une méthode robuste pour recalculer de façon homographique le modèle euclidien d'un objet de la scène sur la reconstruction, on retrouve implicitement la calibration.

## 2.4 Relation affine entre les images

Un des avantages de travailler avec l'espace des disparités réside dans la simplicité à expliciter le lien entre la projection gauche et droite d'un point contenu dans un plan fixé. Dans l'espace rectifié, l'équation du plan support de la face est supposée de la forme suivante :

$$au_g + bu_d + cv + d = 0 \quad (7)$$

En conséquence, un point de l'image droite  $(u_g, v)$  est aussi visible sur l'image gauche en les coordonnées suivantes  $(\frac{au_g + cv + d}{b}, v)$ . La relation est linéaire et la dépendance avec les caractéristiques du plan est particulièrement simple :

$$\begin{pmatrix} u_d \\ v \end{pmatrix} = \begin{pmatrix} \frac{a}{b} & \frac{c}{b} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_g \\ v \end{pmatrix} + \begin{pmatrix} \frac{d}{b} \\ 0 \end{pmatrix}$$

Cette relation est utilisée comme équation de base à nos méthodes de reconstruction géométrique et photométrique présentées dans les sections suivantes, mais elle peut aussi servir à faire de la reconstruction mixte géométrique et photométrique en utilisant l'information de texture des régions [TV95].

## 2.5 Interprétation géométrique

Il est possible d'interpréter la géométrie rectifiée de deux façon différentes. Elle peut être vue comme une projection perspective sur deux images dans un même plan (voir figure 4(a)), ou bien comme une projection orthographique sur chaque image, si elles sont placées perpendiculairement l'une à l'autre (figure 4(b)). Dans le deuxième cas, l'espace entre les images est celui des disparités. Cette double interprétation est la traduction géométrique de l'isomorphisme entre l'espace réel des points visibles de la scène et l'espace des disparités.

On peut remarquer que la déformation qui fait passer de l'espace réel à l'espace des disparités consiste à rejeter les foyers des caméras à l'infini. Il est important de savoir que si l'on travaille avec les coordonnées  $(x, y, z)$ , la projection est perspective, alors que si l'on travaille avec les coordonnées isomorphes  $(u_g, u_d, v)$ , elle est orthographique. L'intérêt de la projection orthographique est que les moments géométriques des images d'une face sont liés aux paramètres de son plan support. Alors, la reconstruction des facettes planes est faite dans l'espace transformé, pour déduire, après transformation inverse (6), l'équation réelle de la facette.

## 2.6 Images intermédiaires

Un des autres avantages de l'utilisation de l'espace des disparités est la facilité qu'il présente pour utiliser des images intermédiaires entre les deux vues. Soit  $\epsilon$  la distance qui repère comme  $\delta$  la position d'une caméra intermédiaire à partir de la position de gauche. Les coordonnées  $(u', v)$  de la projection d'un point sur cette vue intermédiaire sont donc :

$$\left. \begin{aligned} u' &= \frac{x-\epsilon}{z} \\ v &= \frac{y}{z} \end{aligned} \right\}$$

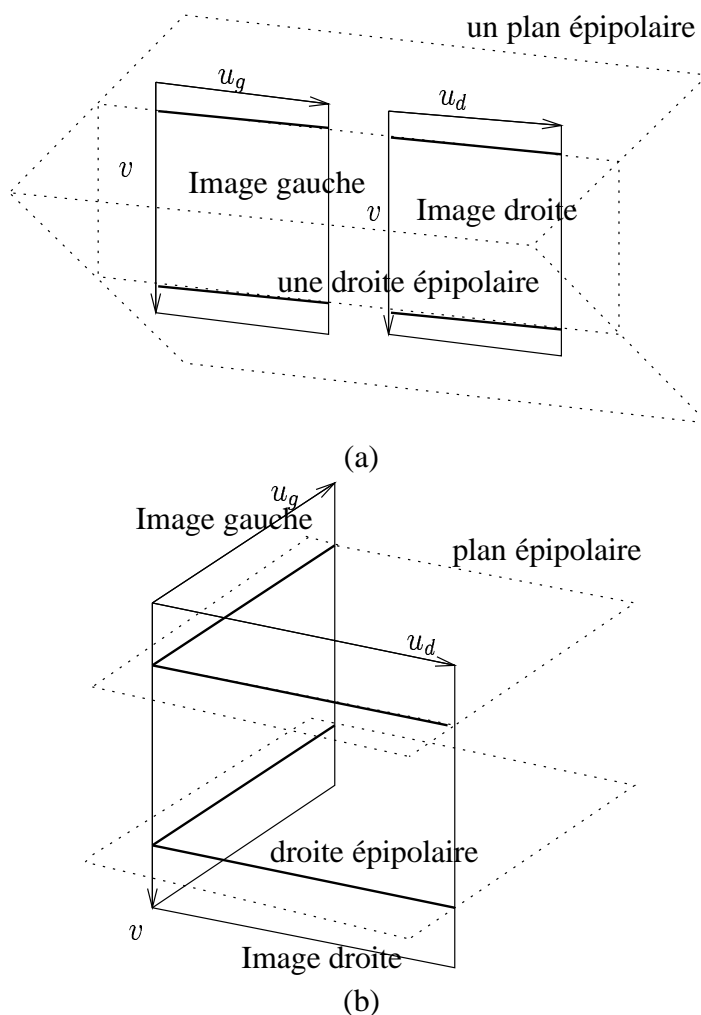


FIG. 4 - (a) *Géométrie des caméras dans le cas rectifié. Les épipolaires sont parallèles.* (b) *Géométrie cohérente à la géométrie du cas (a) après rejet des centres des caméras à l'infini.*

L'équation (6) permet d'éliminer  $X$  et donc d'avoir l'équation dans l'espace des disparités du plan de projection associé à l'image intermédiaire :



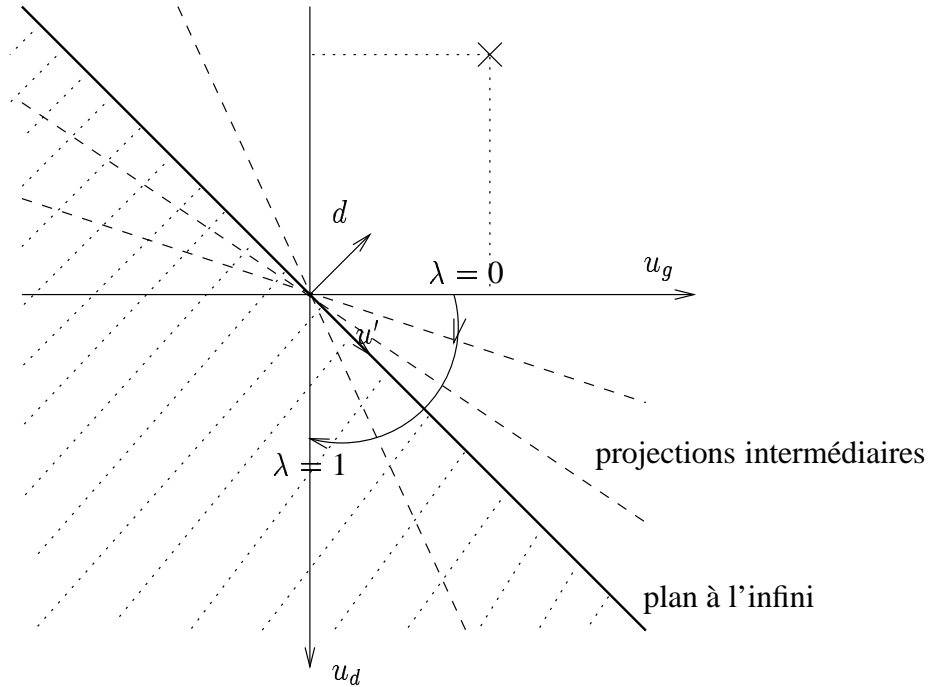


FIG. 5 - Coupe 2D de l'espace des disparités suivant un plan épipolaire (voir figure 4). La partie hachurée n'est pas visible par la caméra.

$$u' = u_g(1 - \lambda) + u_d\lambda \quad \text{avec} \quad \lambda = \frac{\epsilon}{\delta}$$

La coordonnée  $u'$  est donc une simple interpolation des coordonnées  $u_g$  et  $u_d$ . L'image dans la vue intermédiaire d'un point de coordonnées  $(u_g, u_d, v)$  est donc la projection sur un plan passant par l'origine dont la normale est uniquement liée à  $\lambda$ , comme le montre la figure 5.

En choisissant comme vue intermédiaire le milieu des vues gauche et droite, le système de coordonnées  $(u', d)$  défini sur la figure 5 fournit le repère pour représenter des cartes de profondeur où, à chaque point  $(u', v)$ , la disparité  $d$  est donnée.

En conséquence, l'espace des disparités permet d'interpréter une séquence d'image rectifiée comme des projections orthographiques sur un plan en rotation. Cette propriété est très intéressante puisqu'elle permet d'appliquer à des séquences

d'images rectifiées, une méthode de reconstruction volumique qui utilise directement les méthodes de reconstruction d'images scanner développée en imagerie médicale, qui disposent d'architectures spéciales.

## 2.7 Rectification de paires stéréoscopiques

Dans notre méthode de reconstruction, la paire stéréoscopique est rectifiée. Mais, en général, les droites épipolaires ne sont pas parallèles (figure 6). Il est donc nécessaire d'avoir un processus de rectification pour recalculer les nouveaux points de vue, et les nouvelles images qui correspondent à une géométrie épipolaire redressée (figure 6).

Les épipolaires sont parallèles, si et seulement si les deux plans où se forment les images sont confondus, quand les caméras gauche et droite ont la même focale. La rectification consiste donc à choisir un plan 3D comme nouveau plan 'image' commun aux deux caméras. Ce plan est nécessairement parallèle à la droite définie par les deux centres de caméra  $C_g$  et  $C_d$  (figure 2). Il existe une infinité de plans qui conviennent. Ces plans sont tous tangents au cylindre d'axe  $(C_g, C_d)$  et de rayon la focale. Il suffit alors de choisir un plan qui ne produit pas trop de déformations sur les images : celui qui a l'orientation la plus proche de la normale moyenne des deux plans 'image' par exemple [AH88]. Dans le cas de l'utilisation de trois vues, il convient de se reporter à [Aya88]. Ensuite, l'image rectifiée est calculée par reprojection de l'image originale sur le nouveau plan rétinien. Enfin, l'axe de profondeur de la caméra devant être, par convention, confondu avec la normale au plan rétinien, les orientations des deux points de vue sont remises à jour. Les repères sur les images sont choisis alignés avec les épipolaires.

Pour générer des images rectifiées de qualité à partir des images originales, il est nécessaire d'interpoler l'image originale entre ses pixels. La méthode utilisée est une interpolation quadratique à l'intérieur du pixel, les valeurs des quatre sommets étant connues. Cette méthode donne de bons résultats même quand une région de l'image originale est dilatée sur l'image rectifiée.

Le processus de rectification donne des résultats facilement exploitables sur des paires dont la configuration n'est pas trop éloignée de la géométrie d'un couple rectifié (figure 4). Dans le cas contraire, la rectification peut produire des distorsions importantes difficiles à gérer en utilisant de simples images rectangulaires. Néanmoins, à partir de la seule connaissance de l'image originale et de la déformation

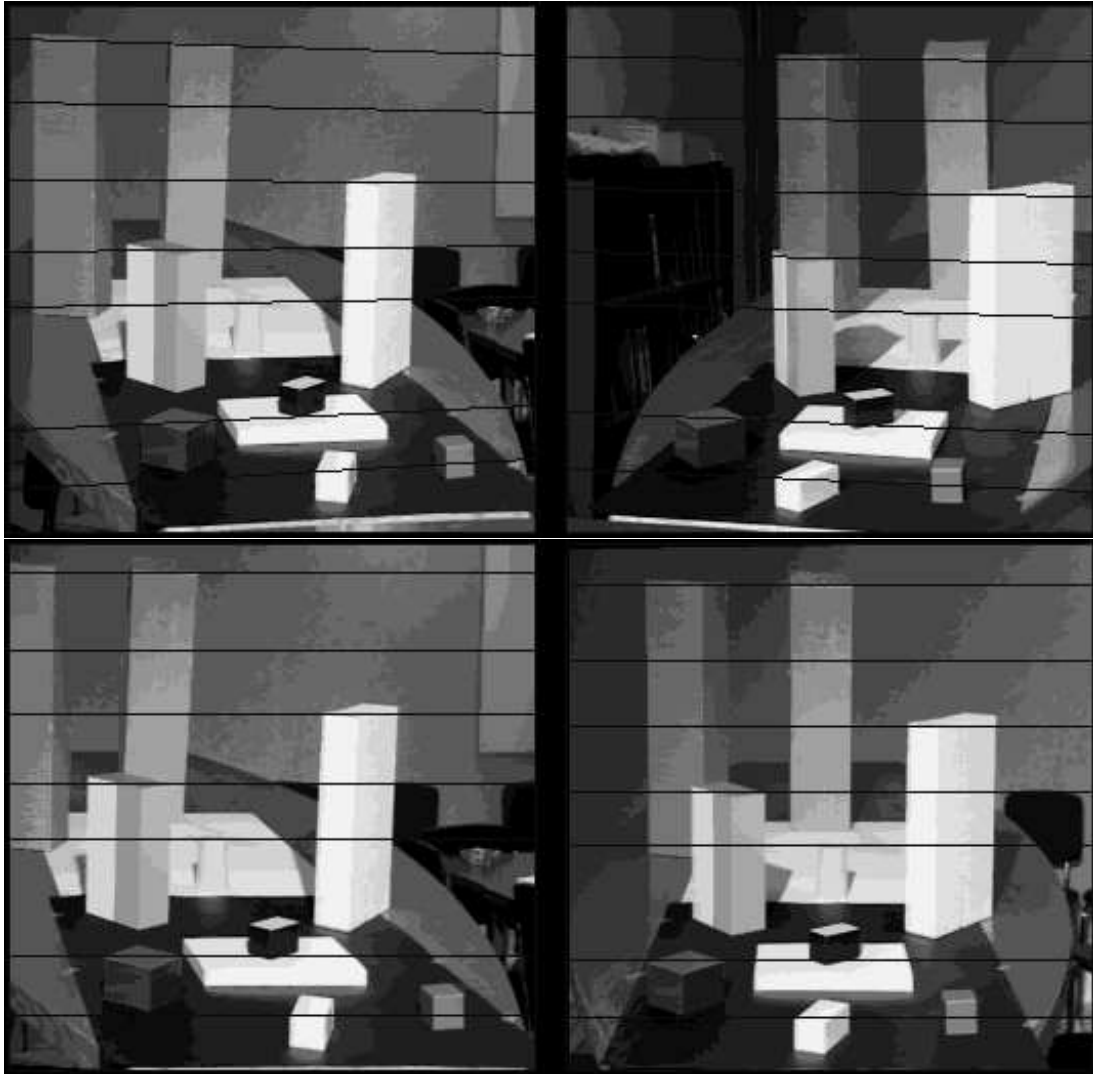


FIG. 6 - *Paire stéréoscopique avant et après rectification. Les épipolaires sont devenues horizontales.*

induite par la rectification, il est possible de faire le calcul des moments des régions sans qu'il soit nécessaire de renumériser l'image rectifiée. Cette dernière approche, même si elle rend le calcul des moments plus compliqué, a l'avantage de ne pas tronquer les informations et peut s'appliquer à toutes les configurations géométriques de caméras.

### 3 La reconstruction de facettes 3D

#### 3.1 La reconstruction géométrique

La linéarité de la projection orthographique permet d'explicitier le lien entre les caractéristiques géométriques d'une facette 3D et de son image en fonction de son positionnement. Soit l'équation du plan, support de la face :

$$au_g + bu_d + cv + d = 0 \quad (8)$$

Sur chaque projection de la face, c'est-à-dire sur la région gauche ou droite de la paire d'images, on calcule les coordonnées du barycentre  $(\overline{u}_g, \overline{v})$  ou  $(\overline{u}_d, \overline{v})$ , ainsi que les trois moments d'inertie  $(\sigma_{u_g, u_g}, \sigma_{u_g, v}, \sigma_{v, v})$  ou  $(\sigma_{u_d, u_d}, \sigma_{u_d, v}, \sigma_{v, v})$ . Ceux-ci sont calculés dans le repère barycentrique de la région. L'utilisation de ces seules données  $(\overline{u}_g, \overline{u}_d, \overline{v})$  et  $(\sigma_{u_g, u_g}, \sigma_{u_g, v}, \sigma_{u_d, u_d}, \sigma_{u_d, v}, \sigma_{v, v})$  s'avère suffisante pour retrouver directement l'équation de la facette. Par intégration simple de (8) sur la face, on obtient une première équation :

$$a\overline{u}_g + b\overline{u}_d + c\overline{v} + d = 0 \quad (9)$$

Le coefficient  $d$  s'exprime donc, grâce à la relation (9), en fonction de la normale  $(a, b, c)$  au plan. On produit trois autres relations, une pour chaque coordonnée, par intégration double de l'équation (8), après soustraction de (9) :

$$a\sigma_{u_g, v} + b\sigma_{u_d, v} + c\sigma_{v, v} = 0 \quad (10)$$

$$a\sigma_{u_g, u_g} + b\sigma_{u_d, u_g} + c\sigma_{v, u_g} = 0 \quad (11)$$

$$a\sigma_{u_g, u_d} + b\sigma_{u_d, u_d} + c\sigma_{v, u_d} = 0 \quad (12)$$

Le coefficient  $d$  n'apparaît pas dans ces dernières équations car les moments d'inertie sont centrés. Les deux dernières expressions (11) et (12) ne sont pas

directement utilisables puisqu'il y apparaît le moment  $\sigma_{u_d, u_g}$  qui n'est pas calculable à partir des images sans avoir la correspondance pixel à pixel entre les deux régions. Toutefois, l'élimination de ce terme entre les deux équations aboutit à l'expression suivante :

$$c = \frac{a^2 \sigma_{u_g, u_g} - b^2 \sigma_{u_g, u_d}}{a \sigma_{v, u_g} - b \sigma_{v, u_d}} \quad (13)$$

De fait, par substitution dans (10), on en déduit le rapport :

$$\frac{a}{b} = \sqrt{\frac{\sigma_{u_d, v}^2 + \sigma_{u_d, u_d} \sigma_{v, v}}{\sigma_{u_g, v}^2 + \sigma_{u_g, u_g} \sigma_{v, v}}} \quad (14)$$

Il existe deux solutions pour ce rapport, l'une positive, l'autre négative. Mais la solution positive est toujours préférée car l'autre correspond à une facette vue d'un côté différent suivant l'une et l'autre des caméras, ce qui est une situation très peu probable dans les scènes courantes. Comme le vecteur  $(a, b, c)$  est normalisé, il est complètement déterminé grâce aux équations (14) et (13). Le plan est donc complètement localisé grâce à l'équation (9). Il est ensuite aisé de faire le passage à la facette reconstruite dans l'espace 3D ordinaire en utilisant (5). L'équation dans l'espace réel de la facette définie par (8) dans l'espace rectifié est en effet d'après (5) :

$$x(a + b) + yc + dz - b\delta = 0$$

Pour que la facette soit reconstruite de manière satisfaisante par cette méthode, elle ne doit pas être occultée. De plus, la reconstruction géométrique donne de mauvais résultats pour des formes de régions dégénérées, les régions très fines par exemple. Mais en dehors de ces deux situations, la reconstruction réalisée est à la fois précise et robuste au bruit (voir section 4 et 5).

### 3.2 Cohérence de la reconstruction géométrique

Quand au cas où la facette à reconstruire est occultée, la méthode précédente va produire une reconstruction plus ou moins biaisée suivant l'ampleur de l'occultation (voir [VT95] pour une l'étude de ce biais d'occultation). Mais, il est possible d'éviter de reconstruire les faces défectueuses en sélectionnant les paires de régions

pertinentes. En effet, les moments d'ordre trois et deux peuvent être utilisés pour vérifier la compatibilité avec l'hypothèse d'une face non occultée. Nous avons montré dans la section 2.4 que la région gauche est la transformée affine de la région droite associée. S'il n'y a pas d'occultation, il existe donc les invariants suivants [FS93] des moments  $M_{ij}$  de chaque région :

$$I_1 = (M_{20}M_{02} - M_{11}^2)/M_{00}^4$$

$$I_2 = (M_{30}^2M_{03}^2 - 6M_{30}M_{21}M_{03} + 4M_{30}M_{12}^3 + 4M_{21}^3M_{03} - 3M_{21}^2M_{12}^2)/M_{00}^{10}$$

$$I_3 = (M_{20}(M_{21}M_{03} - M_{12}^2) - M_{11}(M_{30}M_{03} - M_{21}M_{12}) + M_{02}(M_{30}M_{12} - M_{21}^2))/M_{00}^7$$

Dans le cas où les invariants  $I_1$ ,  $I_2$  ou  $I_3$  sont différents pour les régions gauche et droite, la facette est à rejeter car c'est une facette occultée ou mal segmentée dans la paire. Cette étape de sélection des facettes permet de ne retenir que les facettes qui seront bien reconstruites par la méthode géométrique et de fournir les régions rejetées à un autre algorithme robuste aux occlusions, comme la reconstruction photométrique de la section 3.3, par exemple.

Les moments géométriques peuvent être définis par :

$$M_{ij} = \sum_{(u,v) \in \text{région}} (u - \bar{u})^i (v - \bar{v})^j$$

Mais, il est plus avantageux d'utiliser les moments pondérés par les intensités dans l'image  $I(u, v)$  (éventuellement les intensités sont normalisées) :

$$M_{ij} = \sum_{(u,v) \in \text{région}} I(u, v) (u - \bar{u})^i (v - \bar{v})^j$$

Nous avons étudié sur une scène de synthèse test la robustesse des trois invariants  $I_1$ ,  $I_2$  et  $I_3$ . On observe (figure 7) que les invariants  $I_2$  et  $I_3$  ne sont pas suffisamment stables pour fournir un bon critère. Par contre, l'invariant d'ordre 2,  $I_1$  n'est pas trop sensible aux petites erreurs et néanmoins, il est discriminant (figure 8). Pour éliminer les régions qui ont une occultation supérieure à 10%, il suffit, d'après la figure 8, de sélectionner uniquement les paires de régions dont le rapport des invariants  $I_1$  gauche et droit est entre 0.96 et 1.04.

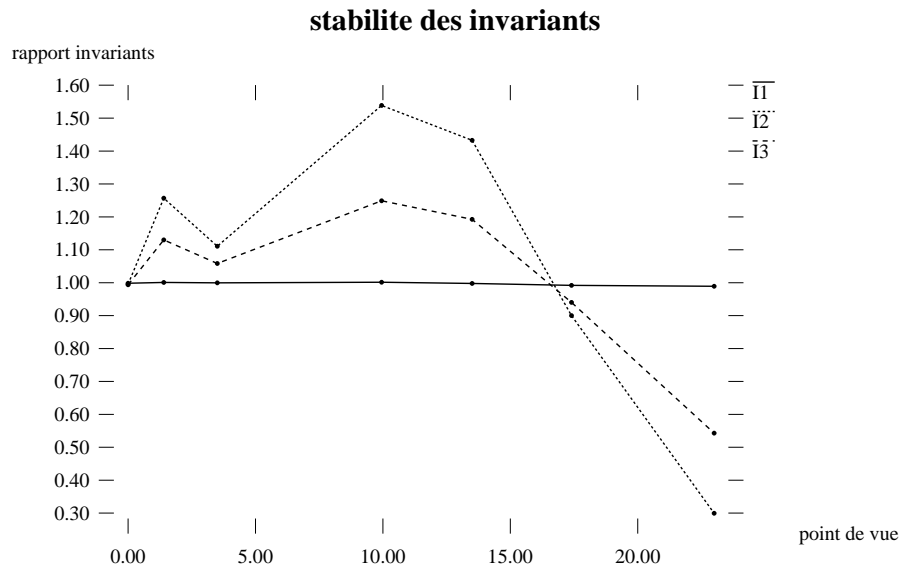


FIG. 7 - *Stabilité de l'invariant  $I_1$  et instabilité des invariants  $I_2$  et  $I_3$  en fonction des changements d'orientation d'une face non occultée.*

### 3.3 La reconstruction photométrique

La reconstruction géométrique utilise la forme des régions gauche et droite. Mais celles-ci ont aussi comme attribut un gradient d'intensité en leur intérieur. Cette information peut servir pour la reconstruction. Il faut remarquer que l'information de gradient a l'avantage de ne pas être perturbée par les occultations.

Soit  $I_g$  et  $I_d$  les distributions d'intensité des régions gauche et droite. On suppose que la distribution est plane. Les surfaces sont donc lambertiennes et les sources lumineuses éloignées.

$$I_g = \alpha_g u_g + \beta_g v + \gamma_g \quad (15)$$

$$I_d = \alpha_d u_d + \beta_d v + \gamma_d \quad (16)$$

On fait alors l'hypothèse habituelle de la reconstruction photométrique, utilisée dans les méthodes créant des cartes de disparités par corrélation [OK85, CM85], où

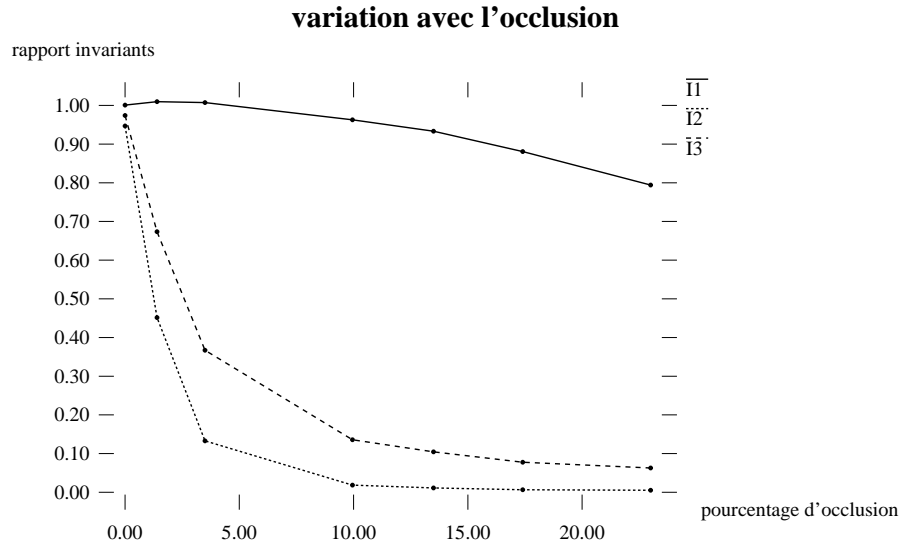


FIG. 8 - Efficacité des invariants  $I_1$ ,  $I_2$  et  $I_3$  en fonction du degré d'occultation.

l'intensité de l'image d'un point est indépendante du point de vue. Cette hypothèse ne prend en compte, ni les variations d'angle solide suivant le point de vue, ni les éventuelles spécularités des surfaces.

L'égalisation des intensités gauche (15) et droite (16) donne directement l'équation de la facette 3D dans l'espace des disparités, ce qui illustre un des autres avantages à travailler dans cet espace :

$$\alpha_g u_g - \alpha_d u_d + (\beta_g - \beta_d)v + \gamma_g - \gamma_d = 0 \quad (17)$$

À partir de (17), il suffit, comme dans le cas géométrique, d'utiliser la transformation inverse (5) pour trouver l'équation de la face dans l'espace réel.

Évidemment, la reconstruction photométrique ne peut pas fonctionner si les gradients d'intensité sont nuls ou trop faibles.



### 3.4 Cohérence de la reconstruction photométrique

Un critère, pour éliminer les régions probablement incorrectes quant à leur photométrie, est disponible comme dans le cas de la méthode géométrique. Ce critère est fondé sur la validité de l'approximation plane de l'intensité de l'image  $I(u, v)$ . Soit l'erreur relative d'approximation plane :

$$\epsilon = \frac{\sum_{(u,v) \in \text{region}} (I(u, v) - \alpha u - \beta v - \gamma)^2}{\alpha^2 + \beta^2}$$

Alors, si ce critère est trop grand, la région n'est pas fiable et sa reconstruction risque d'être erronée.

## 4 Résultats de reconstructions

Les algorithmes de reconstruction ont été testés sur différentes scènes réelles et de synthèse, plus ou moins compliquées. Nous présentons une comparaison des algorithmes géométrique et photométriques ainsi que de leur cadre de validité.

### 4.1 Images de synthèse

Les premiers tests ont été réalisés sur des images synthétiques. Les scènes sont uniquement constituées d'un unique objet. Les images sont segmentées [AMG93] puis mises en correspondance. Ensuite, la reconstruction est comparée au modèle exact utilisé pour générer la scène comme on peut le voir dans la figure 9. On voit sur la vue de côté (b) que la reconstruction (traits en zigzag sur la figure 9) est à la bonne position et à la bonne taille. De même, la géométrie de l'objet est bien respectée, en particulier les angles droits. Le tableau 1 présente les résultats numériques de la comparaison.

### 4.2 Méthodes géométrique et photométrique

La reconstruction a ensuite été testée sur un cube réel. La figure 11 est la paire obtenue suite aux étapes de calibration et de rectification.

Les images de la paire ont été segmentées par un algorithme de croissance de régions [Vin91] puis les régions obtenues ont été mises en correspondance [SVCG89]

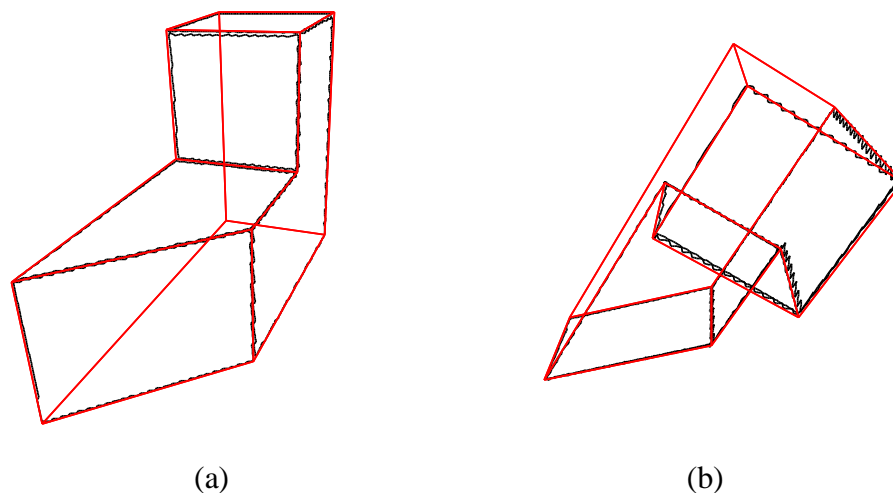


FIG. 9 - Vue de face (a) et de côté (b) de la reconstruction (traits en zigzag) d'un objet de synthèse et du modèle exact associé.

Couple de facette	Angle réel	reconstruit
0-1	135.0	135.45
0-2	90.0	89.61
1-2	90.0	91.24
1-3	90.0	87.79
1-4	90.0	88.98
2-3	90.0	94.33
2-4	0.0	2.45
3-4	90.0	95.20

TAB. 1 - Cohérence des normales reconstruites sur l'objet de synthèse de la figure 9

entre l'image gauche et droite. Les deux méthodes de reconstruction géométrique et photométrique ont été testées. La méthode de reconstruction géométrique (figure 10) donne d'assez bons résultats quant à l'orthogonalité des faces et à leur positionnement relatif.

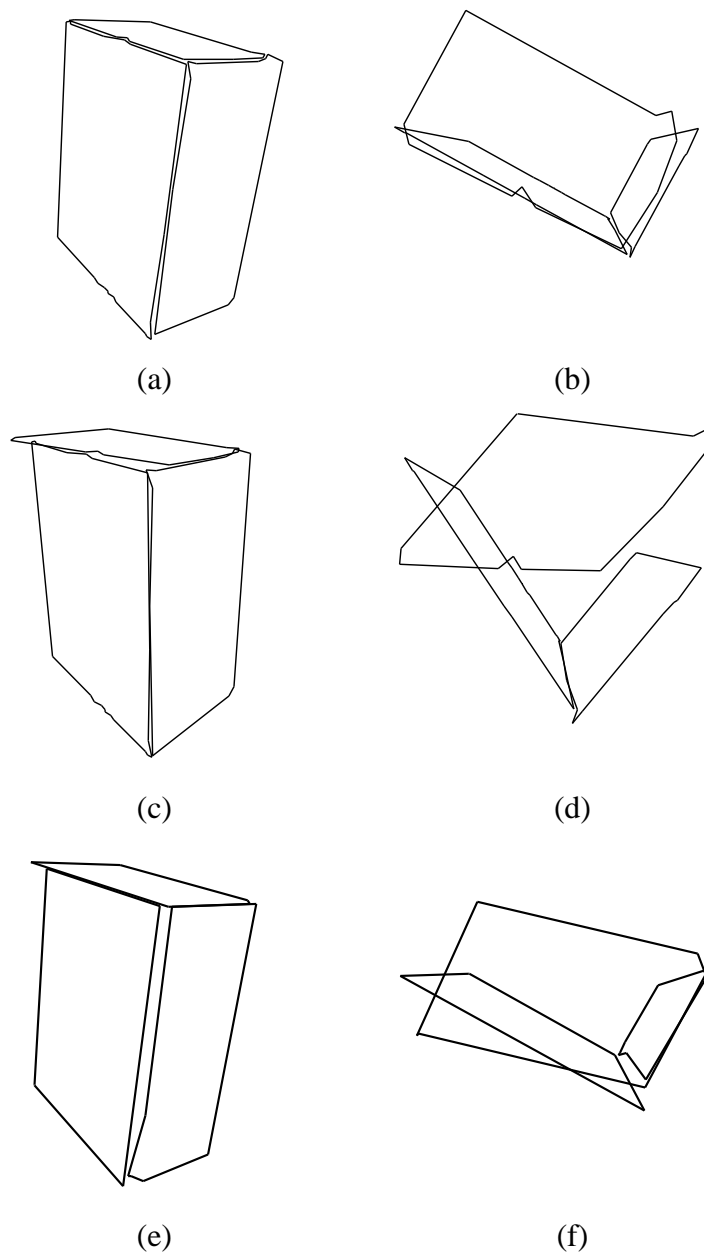


FIG. 10 - *Vue de face (a) et de dessus (b) de la reconstruction du cube de la scène, à partir de la géométrie des régions. Vue de face (c) et de dessus (d) de la reconstruction du cube de la scène, à partir des intensités. Vue de face (e) et de dessus (f) de la reconstruction photométrique, après correction radiométrique des images.*

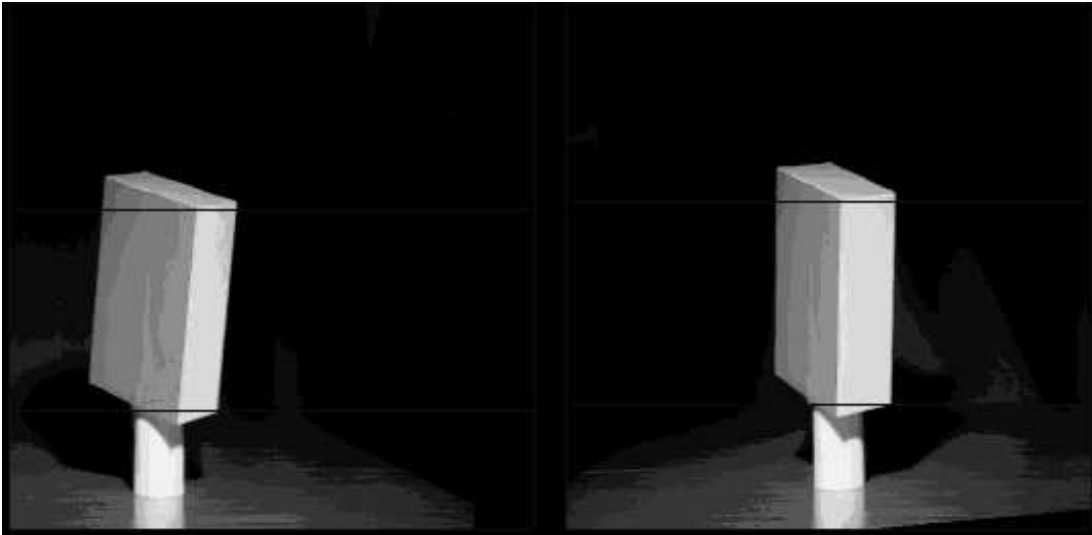


FIG. 11 - Paire stéréoscopique réelle d'un cube après rectification.

La méthode de reconstruction photométrique (figure 10) donne de moins bons résultats car les gradients sur chaque face sont difficiles à estimer précisément. Les raisons de cette difficulté sont de deux types :

- La caméra n'a pas été calibrée radiométriquement. Or, notre caméra atténue les bords de l'image d'une vingtaine de niveaux de gris. Ceci biaise beaucoup la reconstruction photométrique. On peut voir sur le bas de la figure 10, l'amélioration produite par la calibration radiométrique.
- Certaines facettes, en particulier le dessus du cube, ont un gradient très faible qui est mal estimé à cause de la digitalisation. La reconstruction est alors biaisée. La texture et les variations de pigment peuvent aussi produire des erreurs importantes dans l'estimation du gradient.

Au cours de ces tests, l'importance d'une calibration radiométrique de la caméra est apparue. En effet, la calibration radiométrique est nécessaire dès que l'on veut utiliser l'information photométrique de la scène [Tar96].

### 4.3 Sélection des faces

Nous présentons maintenant des résultats sur une scène réelle plus compliquée (figure 6) formée de polyèdres. On voit dans la figure 12 plusieurs parallélépipèdes bien reconstruits. Mais, certaines facettes solitaires semblent flotter dans l'espace sans correspondre à des objets physiques. Les régions utilisées pour les reconstruire sont donc mal segmentées. Ces faces défectueuses sont éliminées en utilisant le test de cohérence fondé sur l'invariant  $I_1$ , comme on le voit sur le bas de la figure 12.

### 4.4 Des objets non polygonaux

La contrainte sur les objets, pour pouvoir les reconstruire, est d'avoir des facettes à contour plan. Ainsi, à partir d'une paire rectifiée (figure 13), segmentée avec un algorithme flou [BSG94], les cercles disposés sur la mire sont reconstruits avec une bonne précision dans un même plan (figure 13). Du fait de la présence d'ellipses, il est difficile de faire la reconstruction avec les techniques classiques de reconstruction qui fond la mise en correspondance des points ou des segments sur les contours.

Notre méthode peut aussi reconstruire une approximation polygonale des objets courbes à la condition que le contour des faces courbes soit plan. En effet, aucune hypothèse n'est faite sur la forme des faces pour la méthode purement géométrique. Dans la figure 14, un ballon, objet sphérique, est reconstruit.

## 5 Comparaison avec d'autres méthodes

Il est important de valider nos algorithmes de reconstruction 3D en calculant l'erreur moyenne de la reconstruction par rapport au modèle de référence s'il est disponible, ou par rapport au résultat d'autres reconstructions. Nous avons donc conçu une méthode de comparaison d'ensembles 3D constitués de points.

### 5.1 Comparaison de données 3D

Dans cette méthode, chaque point 3D du modèle de référence est mis en correspondance avec le point le plus proche en distance de la reconstruction. Une fois la mise en correspondance effectuée, une classification est réalisée sur les déplacements des points en correspondance entre les deux ensembles. Un algorithme

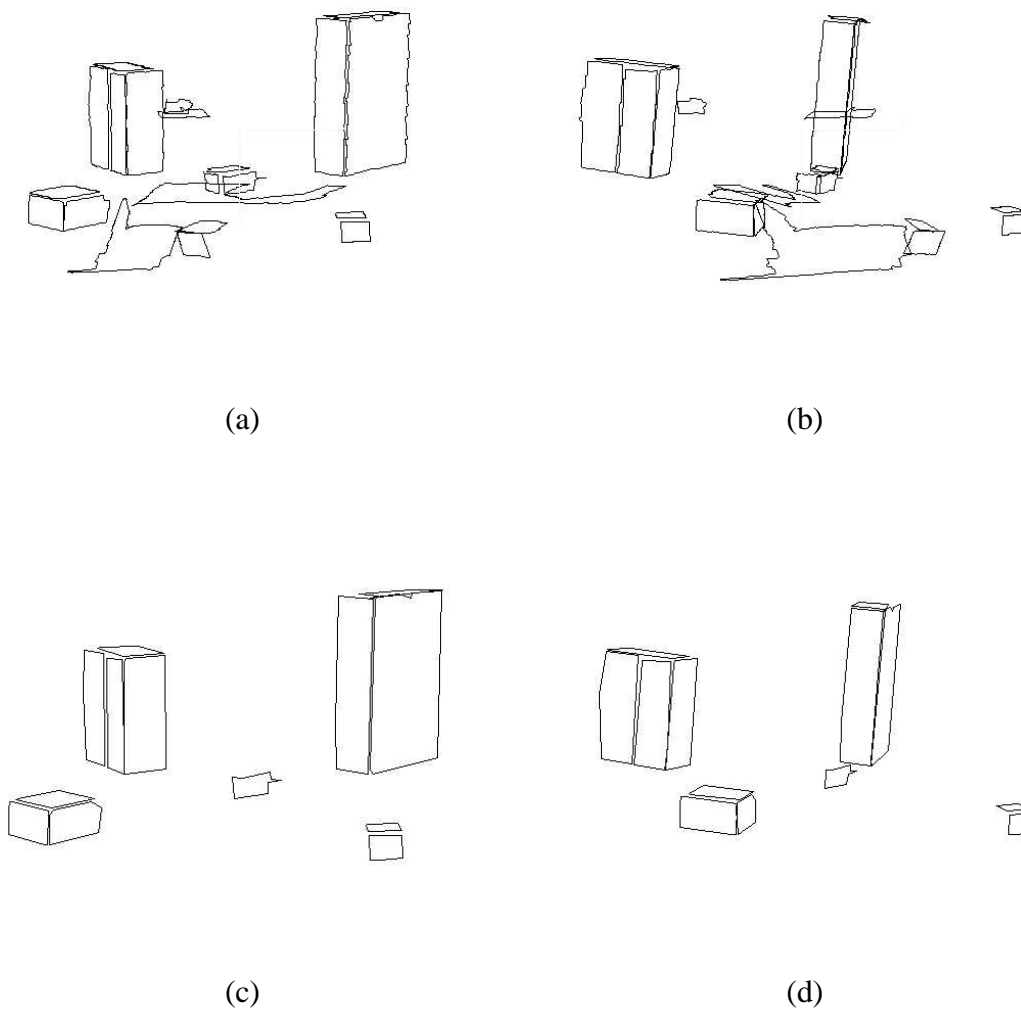


FIG. 12 - Vue de face (a) et de côté (b) de la reconstruction de la scène de la figure 6. Vue de face (c) et de côté (d) de la reconstruction de la même scène après élimination automatique des faces mal segmentées.

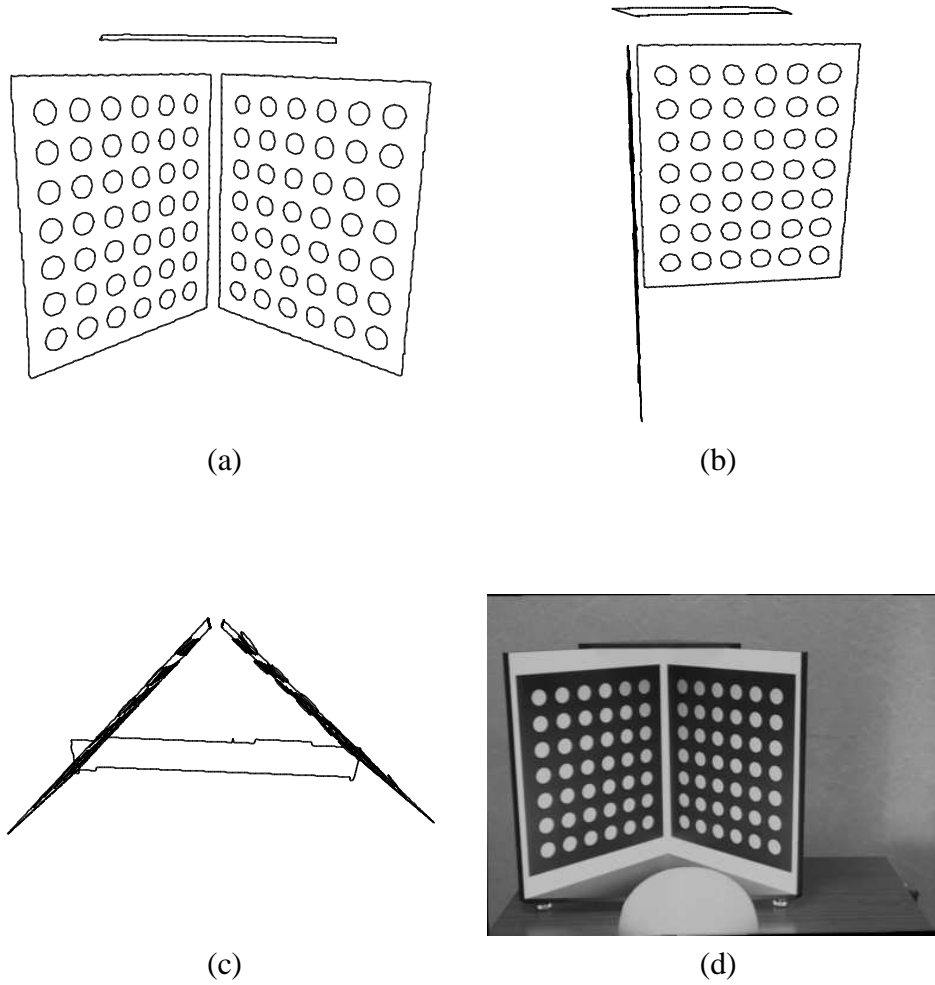


FIG. 13 - *Vue de face (a), de côté (b) et de haut (c) de la reconstruction de la mire de calibration obtenue à partir de l'image (d).*

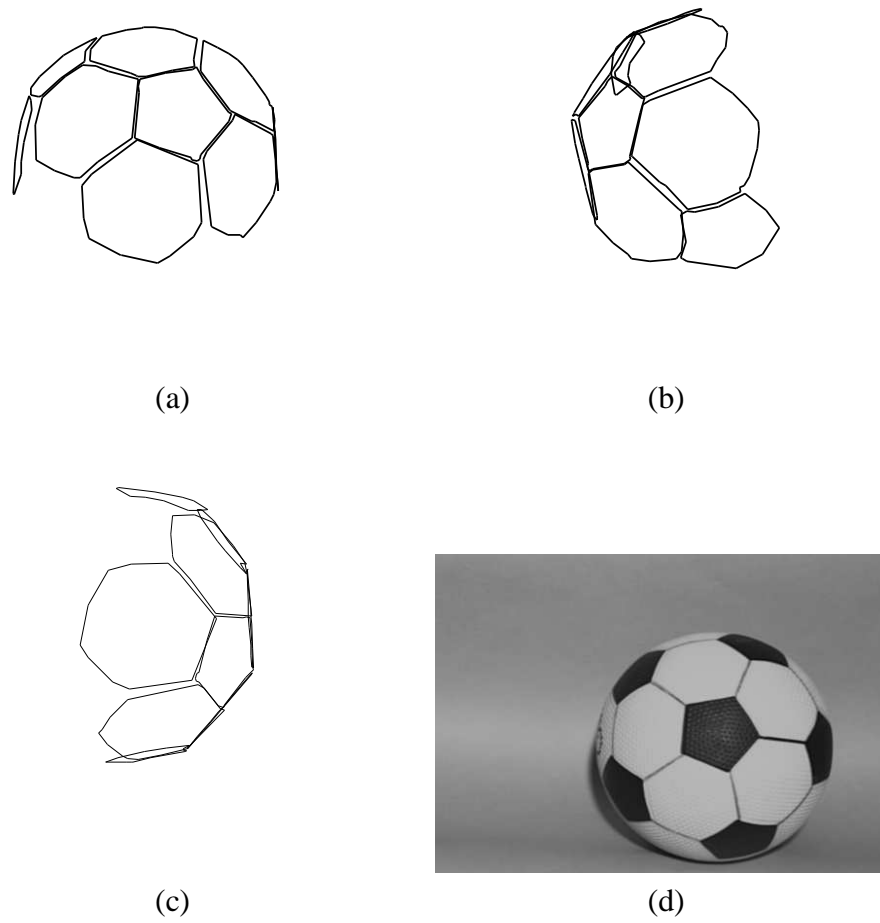


FIG. 14 - *Vue de face (a), de côté (b) et de haut (c) du ballon reconstruit à partir de l'image originale (d).*



de segmentation floue permet de séparer, de manière adaptative, les données en recouvrement de celles qui sont en fausse correspondance. L'algorithme de classification utilisé est celui de Davé [Dav91] avec deux classes : la première récolte les déplacements faibles et la classe de bruit le reste des déplacements 3D. Il est alors aisé de calculer sur la première classe des critères statistiques sur les déplacements, comme le biais et l'écart type moyen des erreurs de reconstruction sur la partie en correspondance.

Cet outil permet l'évaluation objective de nos résultats par rapport à ceux qui furent obtenus par d'autres approches.

## 5.2 Performance de la reconstruction

Voici à présent un exemple des mesures que l'on peut obtenir grâce à notre méthode qui évalue la similarité de deux ensembles 3D. Nous disposons du modèle géométrique exact de la mire que nous utilisons (voir vue (a) de la figure 15). En effet, la mire est l'objet qui a permis de faire la calibration des vues, soit l'estimation de la position et des caractéristiques propres de chaque vue. En conséquence, on dispose, avec une bonne précision [TG95] d'une position de référence de la mire. Celle-ci est présente dans une scène complexe composée de nombreux objets dont un ballon. La reconstruction 3D obtenue par notre méthode est visible dans la vue (b) de la figure 15.

L'espace des déplacements est visible dans la figure 16. Le cœur central de cet ensemble est séparé du reste par une méthode de classification. En effet, les points associés à des déplacements faibles sont selon toute vraisemblance des points communs aux deux ensembles 3D, au bruit et erreurs près.

Une fois la classification réalisée, on visualise sur la figure 17(a) la partie du modèle géométrique de la mire qui a été effectivement reconstruit. Une analyse statistique très simple des erreurs de reconstruction est alors possible. Ces tests nous ont montré que notre reconstruction avait une précision de l'ordre de 4 mm à une distance de 130 cm pour une base de 34 cm (voir tableau 2).

On montre que l'erreur de localisation stéréoscopique en profondeur d'un objet à une distance  $d$  avec une base  $2e$  est donnée par la formule suivante dans le cas rectifié (voir figure 18) :

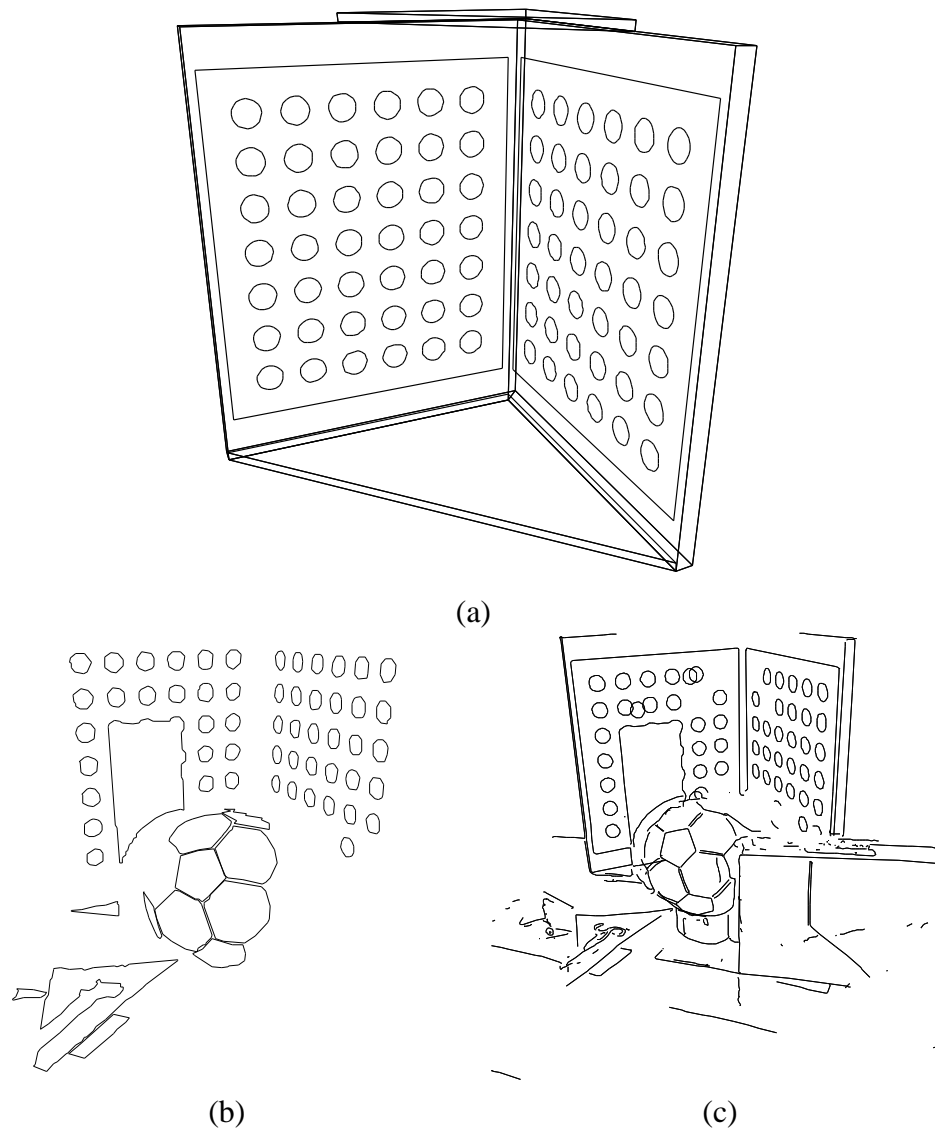


FIG. 15 - *Le modèle de référence (a), notre reconstruction (b) et celle de B. Serra (c).*

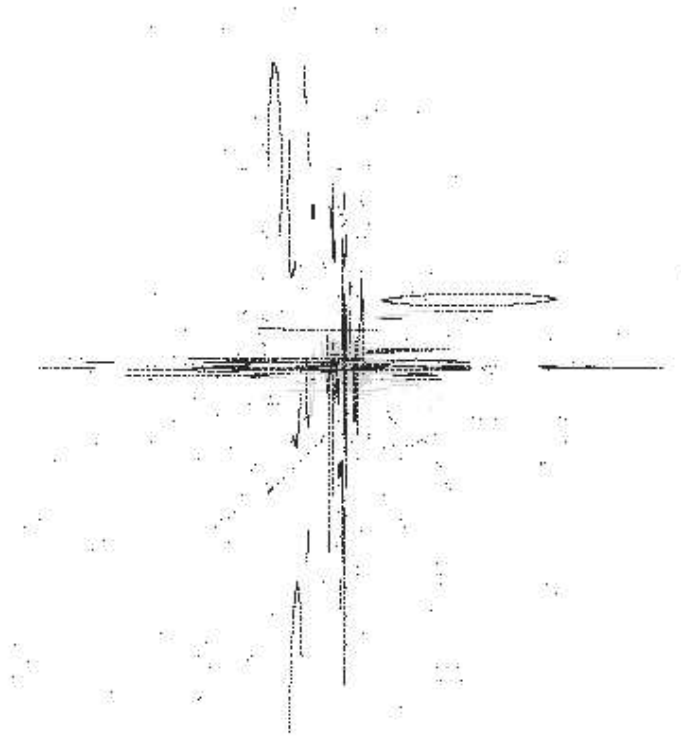


FIG. 16 - *Les déplacements entre le modèle et la reconstruction. La partie centrale, associée à la partie commune entre la reconstruction et le modèle de la mire, est séparée, en gris, de la masse.*

$$\delta = \frac{d^2}{a_u e}$$

où  $a_u$  est l'inverse de la taille d'un pixel de la rétine des caméras. L'erreur théorique est donc de 9 mm en profondeur et 1.2 mm de face. La précision obtenue par notre reconstruction est inférieure à l'erreur induite par la discrétisation de l'image sans explicitement faire une détection sous-pixel des contours des régions. L'observation

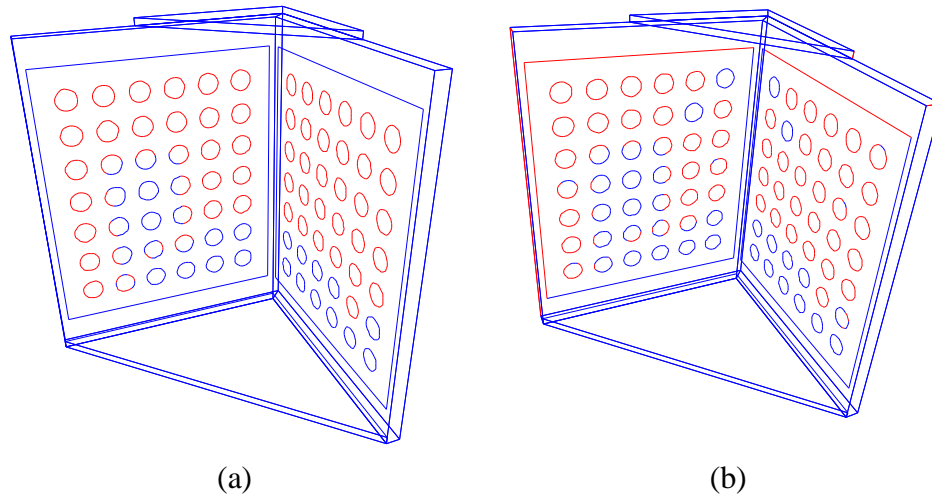


FIG. 17 - (a) Séparation du modèle parfait de la mire en la partie reconstruite et non reconstruite avec notre méthode. (b) Même résultat mais avec la reconstruction de B. Serra.

du tableau 2 montre que l'écart type des erreurs est proche de la limite théorique de 1.2 mm.

### 5.3 Comparaison des méthodes de reconstruction

Notre outil permet en outre de comparer objectivement différentes méthodes de reconstruction (voir figure 15). Ainsi, la méthode de B. Serra [Ber95], de même que la notre, fournissent des résultats sous-pixels en utilisant des primitives de base d'assez haut niveau. En revanche, nous n'utilisons pas comme B. Serra des courbes quelconques mais des contours fermés planaires.

Cette restriction permet d'obtenir une reconstruction, sur cette scène particulière, légèrement plus précise. Ceci se voit lorsque l'on compare les histogrammes des déplacements entre le modèle géométrique de référence et chaque reconstruction de la mire (voir histogramme 19). Cette observation est confirmée par le tableau 2.

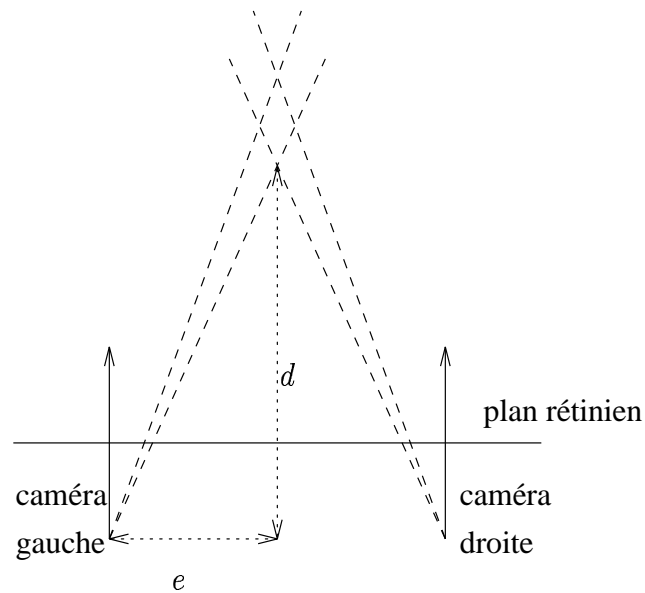


FIG. 18 - *Calcul de l'erreur de localisation stéréoscopique d'un objet à une distance  $d$  avec une base  $2e$ .*

	notre méthode	méthode pastis
taille de la partie sans correspondance	32.5%	38.0%
biais moyen sur partie en correspondance	2.93 mm	3.26 mm
écart type sur partie en correspondance	1.55 mm	1.36 mm

TAB. 2 - *Biais et erreurs moyennes sur la partie en correspondance et taille de la partie sans correspondance entre la mire de référence obtenue par calibration et la reconstruction 3D de la mire.*

Par contre, on remarque sur ce même tableau 2 que notre méthode reconstruit un pourcentage légèrement plus faible de la mire. Cela s'explique par le fait que nous nous sommes restreint à des courbes fermées comme on peut l'observer en comparant les figures 17.

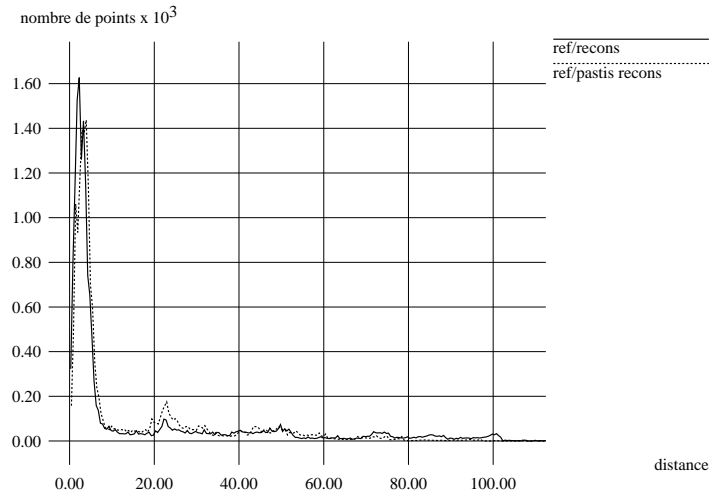


FIG. 19 - *Histogrammes des déplacements entre le modèle et les reconstructions obtenues avec notre méthode et celle de B. Serra.*

## 6 conclusion

Notre méthode de reconstruction globale de facettes 3D est exacte. Le fait qu'elle utilise des images rectifiées n'altère en rien ces performances puisqu'à partir d'un couple quelconque d'images, il est possible de recalculer une paire rectifiée, quand les deux caméras sont calibrées. Les méthodes proposées, fondées sur la géométrie ou sur la photométrie, ont été testées avec succès sur différentes scènes réelles. Elles sont robustes au bruit et précises puisqu'elles permettent d'atteindre une précision équivalente aux résultats d'une méthode sous-pixel. De plus, leur combinaison permet de traiter les cas de facettes occultées.

De surcroît, pour chaque méthode, nous avons tiré profit de la simplicité à formuler les problèmes de reconstruction de facettes, lorsqu'on travaille dans l'*espace des disparités* explicité dans ce rapport.

Notre reconstruction utilise la primitive face 3D. Mais, chaque primitive ayant ses avantages et ses limites, une perspective intéressante à l'avenir, nous semble de

combiner les facettes planes avec les contours pour obtenir la reconstruction la plus complète et la plus robuste possible.

## Références

- [AH88] Ayache (N.) et Hansen (C.). – Rectification of images for binocular and trinocular stereovision. *In: Ninth International Conference on Pattern Recognition (Rome, Italy, November 14–17, 1988)*. pp. 11–16. – Washington, DC, 1988.
- [Alo90] Aloimonos (John Y.). – Perspective approximations. *Image and Vision Computing*, vol. 8, n° 3, August 1990, pp. 179–192.
- [AMG93] Ackah-Miezan (A.) et Gagalowicz (A.). – Discrete models for energy minimizing segmentation. *In: Proceedings of the 3rd International Conference on Computer Vision*, pp. 200–207. – 11-13 May, Berlin, 1993. <http://www-rocq.inria.fr/syntim/textes/iccv93-eng.html>.
- [Aya88] Ayache (N.). – *Construction et fusion de cartes visuelles 3D: Application à la robotique mobile*. – Thèse de PhD, Université Paris-XI, 1988.
- [Ber95] Berthod (B. Serraand M.). – Optimal subpixel matching of contour chains and segments. *In: Proc. ICCV 1995*. – Cambridge, MA, 1995.
- [BMV93] Boufama (B.), Mohr (R.) et Veillon (F.). – Euclidean constraints for uncalibrated reconstruction. *In: IEEE International Conference on Computer Vision*, pp. 466–470. – Berlin, March 1993.
- [BSG94] Boujemaa (N.), Stamon (G.) et Gagalowicz (A.). – Modélisation floue pour la segmentation d’images. *In: 9ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*.
- [CM85] Cochran (S.D.) et Medioni (G.). – 3-d surface description from binocular stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, n° 2, 1985, pp. 139–198.

- [CM94] Chabbi (H.) et Masini (G.). – Construction de facettes 3D à partir de triplets stéréoscopiques en utilisant les invariants projectifs. *In : 9ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle.*
- [CO85] Cyganski (D.) et Orr (J.A.). – Applications of tensor theory to object recognition and orientation determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, n° 6, 1985, pp. 662–674.
- [Dav91] Dave (R.N.). – Characterization and detection of noise in clustering. *Pattern Recognition Letters*, vol. 12, 1991, pp. 657–664.
- [FS93] Flusser (J.) et Suk (T.). – Pattern recognition by affine moment invariants. *Pattern Recognition*, vol. 26, 1993, pp. 167–174.
- [JF92] Jourdan (G.) et Feldmar (J.). – *Reconstruction géométrique et reconnaissance de modèles.* – Rapport de dea, INRIA, 1992.
- [Lus87] Lustman (F.). – *Vision stéréoscopique et perception du mouvement en vision artificielle.* – Thèse de PhD, Université Paris-Sud, 1987.
- [LVG92] Lutton (E.), Vézien (J-M.) et Gagalowicz (A.). – Model based stereo reconstruction by energy minimization. *In : Proceedings, ISCIS.* – Bordeaux, France, 1992.
- [Ma93] Ma (S. De). – Conics-based stereo, motion estimation and pose determination. *International Journal of Computer Vision*, vol. 10, n° 1, 1993.
- [MF92] Maybank (S.J.) et Faugeras (O.D.). – A theory of self-calibration of a moving camera. *IJCV*, vol. 8, n° 2, 1992, pp. 123–151.
- [OK85] Ohta (Y.) et Kanade (T.). – Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, n° 2, 1985, pp. 139–198.
- [Ran92] Randriamasy (S.). – *Segmentation descendante coopérative en régions de paires d'images stéréoscopiques.* – Thèse de PhD, Université Paris-IX Dauphine, 1992.



- [RZMF92] Rothwell (C.A.), Zisserman (A.), Marinos (C.I.) et Forsyth (D.A.). – Relative motion and pose from arbitrary plane curves. *Image and Vision Computing*, vol. 10, n° 4, 1992, pp. 250–262.
- [SVCG89] Sander (P.), Vinet (L.), Cohen (L.) et Gagalowicz (A.). – Hierarchical regions based stereo matching. In : *Proceedings of the Sixth Scandinavian Conference on Image Analysis*, pp. 71–78. – Oulu, Finland, June 1989.
- [Tar96] Tarel (J.P.). – Une méthode de calibration radiométrique de caméra à focale variable. In : *10ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*. – Rennes, France, 1996.
- [TG95] Tarel (J.P.) et Gagalowicz (A.). – Calibration de caméra à base d’ellipses. *Traitement du Signal*, vol. 12, n° 2, 1995, pp. 177–187. – <http://www-rocq.inria.fr/syntim/textes/calib-eng.html>.
- [TV95] Tarel (J.-P.) et Vézien (J.-M.). – A generic approach for planar patches stereo reconstruction. In : *Proceedings of the Scandinavian Conference on Image Analysis*, pp. 1061–1070. – Uppsala, Sweden, 1995. <http://www-rocq.inria.fr/syntim/textes/scia95-eng.html>.
- [Véz95] Vézien (J.M.). – *Techniques de reconstruction globale par analyse de paires d’images stéréoscopiques*. – Thèse de PhD, Université Paris-VII, 1995.
- [Vin91] Vinet (L.). – *Segmentation et mise en correspondance de régions de paires d’images stéréoscopiques*. – Thèse de PhD, Université Paris-IX Dauphine, 1991.
- [VT95] Vézien (J.M.) et Tarel (J.P.). – *A Generic Approach for Planar Patches Stereo Reconstruction*. – Research report 2507, INRIA, 1995. <http://www-rocq.inria.fr/syntim/textes/RR-2507-eng.html>.



---

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,  
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY  
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex  
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1  
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex  
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

---

Éditeur  
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)  
ISSN 0249-6399