



Une approche floue du recalage 3D : généricité et robustesse

Jean-Philippe Tarel, Nozha Boujemaa

► **To cite this version:**

Jean-Philippe Tarel, Nozha Boujemaa. Une approche floue du recalage 3D : généricité et robustesse. [Rapport de recherche] RR-2716, INRIA. 1995, pp.31. <inria-00073976>

HAL Id: inria-00073976

<https://hal.inria.fr/inria-00073976>

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.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

***Une approche floue du recalage 3D : généralité
et robustesse***

Jean-Philippe Tarel, Nozha Boujemaa

N° 2716

Novembre 1995

PROGRAMME 4



***rapport
de recherche***

Une approche floue du recalage 3D : généricité et robustesse

Jean-Philippe Tarel *, Nozha Boujemaa **

Programme 4 — Robotique, image et vision
Projet Syntim

Rapport de recherche n° 2716 — Novembre 1995 — 31 pages

Résumé : Le *recalage* 3D est un problème important en vision par ordinateur qui a de nombreuses applications, tant pour obtenir une reconstruction incrémentale complète à partir de primitives 3D extraites de différents points de vue, que pour interpréter une scène reconstruite à partir du modèle des objets qui la composent. Nous nous intéressons ici au recalage rigide entre deux ensembles 3D qui présentent une partie commune. Notre approche est fondée sur l'intégration d'une méthode de classification floue adaptée à un contexte fortement bruité, pour obtenir le schéma général d'une méthode de recalage robuste. Quelques résultats sont décrits pour illustrer les avantages de la méthode, qui peut notamment gérer, de manière unifiée, plusieurs solutions recalées. Enfin, nous présentons un outil pour comparer deux ensembles 3D et l'utilisons pour valider notre méthode.

Mots-clé : Recalage 3D, Mise en correspondance, Classification floue avec bruit, Interprétation avec modèles.

(Abstract: *pto*)

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

** E-mail : Nozha.Boujemaa@inria.fr, LI / École d'Ingénieurs en Informatique pour l'Industrie (E3i), 64 Avenue Jean Portalis-Technopole boîte 4, 37913 Tours Cedex 9, France.

Fuzzy approach for a Generic and Robust 3D Registration method

Abstract: An important problem in computer vision is to recover how features extracted from images are connected to existing models. In this paper, we focus on solving the *registration* problem, i.e obtaining rigid displacement parameters between several 3D data sets, whether partial or exhaustive. The difficulty of this problem is to obtain a method which is robust to outliers and at the same time accurate. We present a general method performing robust 3D location and fitting based on fuzzy clustering method. The fuzzy set approach is known to its practical efficiency in uncertain environment. To illustrate the advantages of this approach on the registration problem, we show results on synthetic and real 3D data. Moreover, we propose a tool to compare two 3D data sets in terms of similarity and dissemblance.

Key-words: 3D registration, matching, fuzzy clustering with noise, model-based interpretation.

1 Introduction

L'interprétation de données 3D à l'aide de modèles et la concaténation de reconstructions 3D obtenues sous différents angles de vue, sont des problèmes qui se ramènent à des recalages 3D. En effet, le recalage est un sujet qui suscite un intérêt croissant de par ses applications dans de nombreux domaines aussi différents que l'imagerie médicale tridimensionnelle [Aya93], l'imagerie aérienne [HH94] ou la cartographie [KPLR91].

Les méthodes de recalage peuvent se classer en deux catégories principales : avec recherche de mises en correspondance combinatoires et avec appariements explicites itérés.

Le premier type de méthodes s'appuie sur la géométrie locale pour construire des hypothèses valides de mises en correspondance. La difficulté est alors de sélectionner l'hypothèse qui se répète le plus. Celle-ci correspond à un déplacement qui recalcule une partie commune aux deux ensembles de données 3D. On peut classer dans cette catégorie les méthodes de recalage fondées sur la transformation de Hough et les tables de hachage [FH83, Aya93, RH91, GA92]. Or celles-ci ont souvent le désavantage d'être facilement perturbées par de faux appariements lors de l'estimation du déplacement.

La deuxième classe de méthodes réalise l'étape de mise en correspondance et de recalage de manière itérative. Puisque l'appariement est remis en cause à chaque étape, il n'est pas gênant que la mise en correspondance ne soit qu'approximative dans les premières étapes de l'algorithme. Un critère explicite de plus proche voisin convient bien pour faire ces appariements. Néanmoins, l'appariement initial ne doit pas être trop perturbé par les faux appariements, au risque de converger vers une solution absurde. Ces méthodes nécessitent donc que le déplacement entre les deux jeux de données ne soit pas trop important. Cette contrainte limite la généralité de ce type de méthodes, mais les recalages obtenus sont généralement très précis. On peut citer dans cette catégorie les méthodes basées sur l'approche "Iterative Closest Point" (ICP) [BM92, Zha94, KVB93].

Dans ce rapport, nous proposons de combiner les deux types de méthodes précédemment décrits pour tirer profit de leurs avantages complémentaires : la possibilité de grands déplacements avec une méthode combinatoire et la précision avec une méthode itérative.

De manière générale, nous pouvons distinguer trois critères de robustesse pour une méthode qui procède par minimisation:

- la robustesse au bruit sur les données souvent modélisé par une variable aléatoire Gaussienne centrée,
- la robustesse aux points aberrants qui sont des données parasites non significatives,
- et la robustesse à l'initialisation.

Dans la section 2, nous présentons un schéma en trois étapes d'une méthode robuste au bruit, aux points aberrants et à l'initialisation, pour faire le recalage 3D. La première étape consiste à créer l'ensemble des déplacements qui peuvent superposer une partie des deux ensembles 3D (section 2.1). Nous introduisons un premier filtrage des appariements à l'aide d'un coefficient de similarité associé à chacun pour que seules des primitives assez semblables soient considérées. La deuxième étape consiste à déterminer les déplacements significatifs dans cet ensemble (section 2.2). Nous avons, pour cela, développé un algorithme de classification floue fondée sur la méthode de Davé [Dav91] ayant la particularité de gérer les données bruitées et les points parasites. Ceci nous a permis d'approcher de manière robuste les solutions optimales. A ce stade, une dernière étape est nécessaire pour réaliser un recalage fin et atteindre un résultat plus précis (section 2.3). Dans la section 3, les possibilités de la méthode sont illustrées sur quelques exemples de jeux de données réelles. Enfin, dans la dernière section, un outil de comparaison et de validation des méthodes de recalage est proposé et utilisé sur nos résultats.

2 Schéma général du recalage 3D avec une approche floue

La méthode réalise les trois étapes suivantes (figure 1): un recalage approché mais robuste par une approche floue après une phase d'appariement combinatoire, puis un raffinement itératif du recalage par une méthode de mise en correspondance explicite.

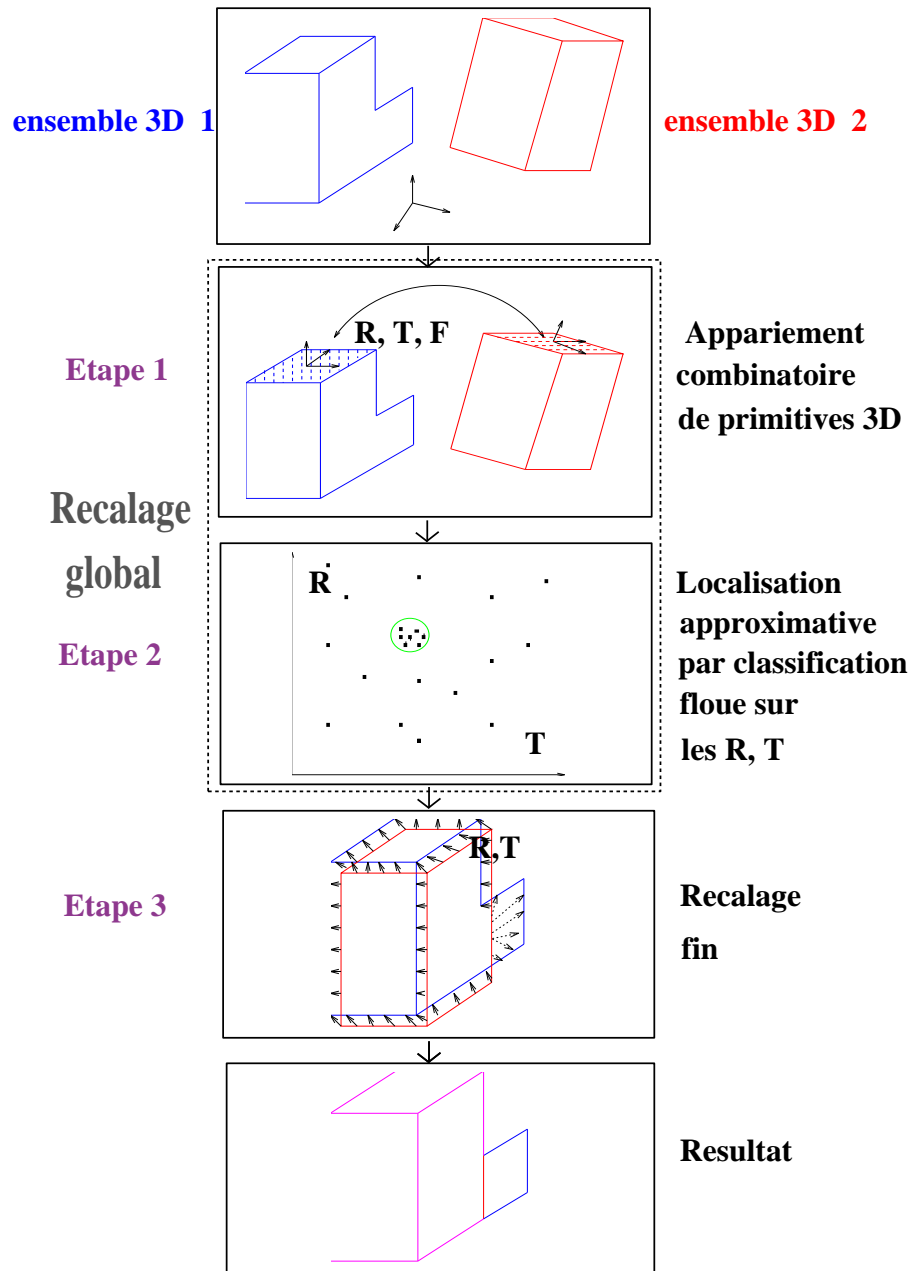


FIG. 1 - Approche générale du recalage 3D flou.

2.1 Mise en correspondance de primitives 3D

La première étape de la méthode consiste à faire la mise en correspondance 3D de manière à explorer toutes les possibilités. Un ensemble 3D est constitué d'entités géométriques comme des points, segments, courbes ou facettes. Il faut choisir la primitive de base à appairier, de façon à permettre les deux estimations suivantes entre deux primitives quelconques :

- calcul des déplacements qui les superposent,
- calcul d'une valeur de ressemblance entre les deux primitives.

Dans cet article, nous avons utilisé des facettes, mais de nombreuses autres primitives sont aussi utilisables, selon le type des données manipulées : base 3D de 4 points [RH91], couple de segments, ou repère local de Frénet sur des courbes [GA92].

Une mesure importante de la combinatoire du problème d'appariement est le rapport SNR du nombre de mises en correspondance justes sur leur nombre total. Par exemple, si l'on suppose que le nombre de primitives dans l'ensemble à recalcr est N_{prim} et que l'autre ensemble lui est identique à un déplacement près, le nombre total d'appariements possibles est $N_{prim}^2 N_{depla}$, alors qu'il n'y a que N_{prim} appariements justes. Le rapport SNR est donc dans ce cas égal à $\frac{1}{N_{prim} N_{depla}}$. Or, ce rapport décroît rapidement avec le nombre de primitives N_{prim} . Il est donc préférable de choisir des primitives relativement abstraites, qui sont en petit nombre comme les facettes, pour limiter la combinatoire.

Détaillons à présent notre choix pour faire le calcul des déplacements et du facteur de ressemblance dans le cas d'une paire de faces. Une facette peut être partiellement caractérisée par son barycentre et ses axes principaux d'inertie. Ce point et les axes normalisés fournissent un repère propre à la face. Il devient alors possible de calculer les déplacements qui superposent deux faces quelconques à partir de ces repères propres (figure 2). Il y a plusieurs déplacements possibles pour un même couple de facettes, mais ce nombre est fini dans les cas non dégénérés. Soit N_{depla} ce nombre. Par exemple, pour des rectangles, $N_{depla} = 4$, alors que pour des carrés $N_{depla} = 8$. Le cas dégénéré est celui où la face est circulaire, puisque les axes propres sont connus à une rotation près.

De plus, on obtient par ce calcul deux valeurs intrinsèques à la face i qui sont la longueur des deux plus longs axes principaux d'inertie α_i et β_i . À partir de ces données invariantes par déplacement rigide, il est possible d'affecter à chaque

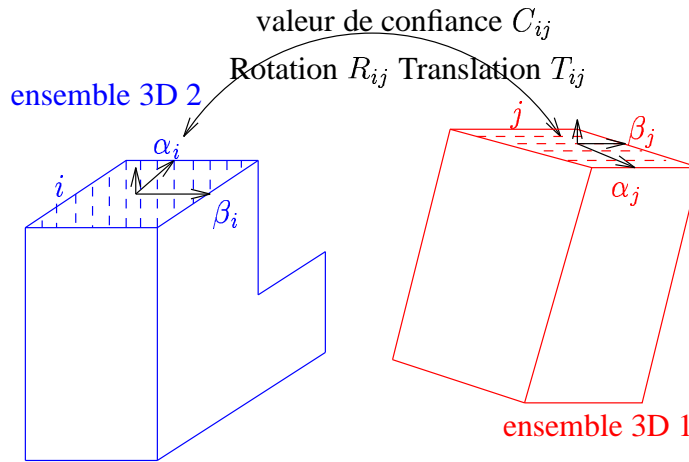


FIG. 2 - Appariement des faces i et j de deux ensembles de faces. Pour une paire (i, j) de rectangles. Quatre déplacements (R_{ij}, T_{ij}) sont au moins possibles.

appariement une valeur de confiance qui décrit la qualité de cet appariement. Dans notre exemple, nous avons choisi de combiner les valeurs des axes principaux d'inertie, pour un couple de faces (i, j) , de la manière suivante :

$$C_{ij} = \min\left(1, \frac{I}{|\alpha_i - \alpha_j|}\right) \min\left(1, \frac{I}{|\beta_i - \beta_j|}\right) \quad (1)$$

et ainsi obtenir une valeur entre 0 et 1 qui mesure leur ressemblance. Le paramètre I est fixé à partir de l'estimation de l'erreur moyenne des données 3D (section 4.2). Plus les deux facettes sont de géométries différentes, et plus la valeur du critère est proche de 0.

La rapidité de la décroissance avec l'écart des longueurs des axes est réglée par le coefficient I (voir figure 3). Les appariements qui ont une valeur de similarité très faible (entre 0 et 0.1) sont ignorés dans la suite car ils ne présentent pas d'intérêt et cela évite l'explosion combinatoire.

En conséquence, après l'étape de mise en correspondance, on dispose d'un ensemble d'appariements de facettes, où sont associés à chaque paire un facteur de ressemblance et des déplacements.

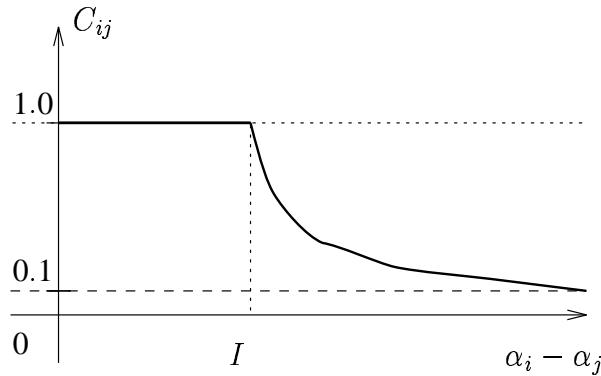


FIG. 3 - Variations de la valeur de confiance C_{ij} avec la différence de la longueur du premier axe propre $\alpha_i - \alpha_j$. I est l'erreur moyenne sur les données 3D.

2.2 Localisation 3D approchée

Dans la première étape, un ensemble d'appariements a été obtenu, où chaque élément est associé à des déplacements et un facteur de confiance. Un déplacement 3D peut être représenté de diverses manières comme un point x de \mathbb{R}^6 , avec trois termes de translation et les trois angles d'Euler, par exemple. Nous extrayons alors l'ensemble des déplacements 3D qui superposent au moins une facette. On réalise donc le changement d'indice suivant :

$$(R_{ij}, T_{ij}) \rightarrow x_{k(i,j)} \quad (2)$$

$$C_{ij} \rightarrow C_{k(i,j)}$$

Dans cet ensemble $(x_k)_k$ de déplacements, *un point d'accumulation correspond à une hypothèse intéressante de recalage* entre l'une des parties commune aux deux ensembles 3D (figure 4). Comme les méthodes de classification floue sont particulièrement robustes, elles sont appropriées pour sélectionner ces points d'accumulation. Dans ce cadre, chaque point d'accumulation correspond au prototype d'une classe. Puisque l'algorithme de classification manipule en même temps plusieurs classes, il gère donc plusieurs hypothèses de déplacement en parallèle, en essayant de recouvrir au maximum l'ensemble des déplacements possibles. Un deuxième avantage

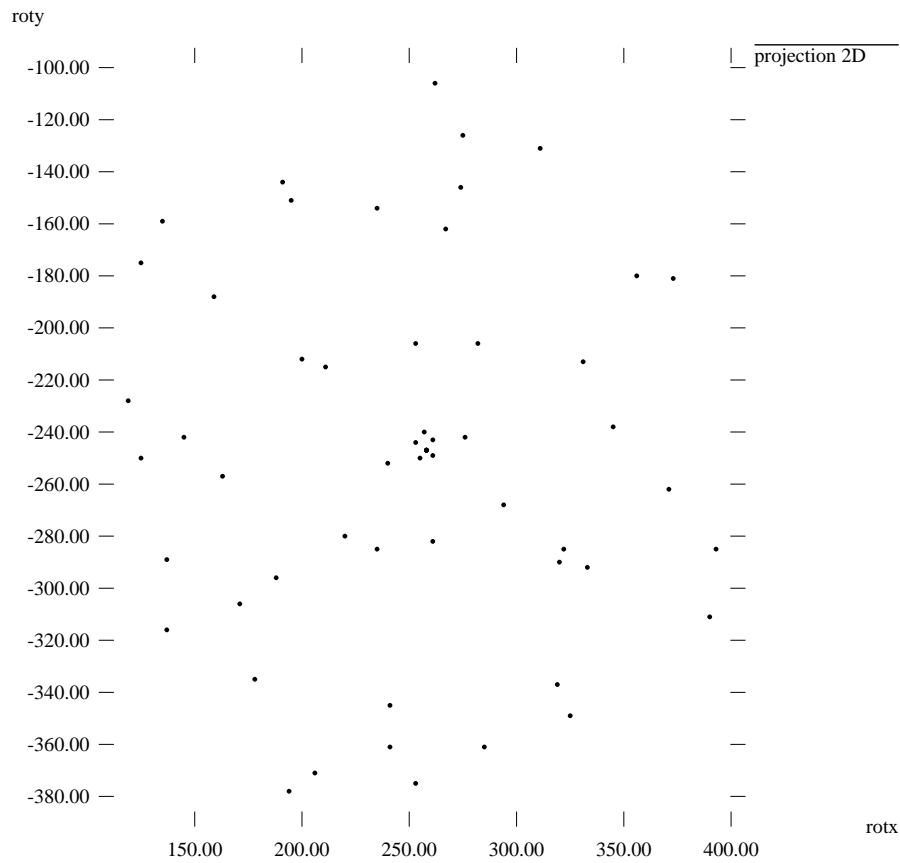


FIG. 4 - Projection 2D de l'ensemble des rotations 3D à classifier pour obtenir le point d'accumulation central correspondant à une hypothèse intéressante pour le recalage.

d'utiliser une méthode de classification est de ne pas être pénalisé par la discrétisation de l'espace d'accumulation comme dans la transformée de Hough [Sto87].

2.2.1 L'approche floue

La modélisation des problèmes de vision par ordinateur avec des sous-ensembles flous [Bou93] permet de reporter les prises de décision en phase ultime d'un processus de traitement, minimisant ainsi la propagation des erreurs décisionnelles dans des phases intermédiaires. Des approches locales ou globales peuvent être utilisées pour définir les fonctions d'appartenance floues caractéristiques des propriétés des données traitées en vision. Le premier type d'approche, proposé par Pal et Rosenfeld, généralise les concepts de topologie et de géométrie (connexité, adjacence, compacité) dans le contexte des sous ensembles flous. Les approches globales sont fondées sur les techniques de classification adaptative floue introduites initialement par Ruspini, Dunn et Bezdek, qui minimisent une famille de fonctionnelles de base : les C-Moyennes Floues (FCM) [Bez81]. Le principe de la classification floue est d'autoriser une appartenance partielle ou distribuée d'une forme donnée à l'ensemble des classes de l'espace des caractéristiques. La classification floue, en y associant une décision dépendante du contexte spatiale, s'est révélée très efficace pour la segmentation d'images. Dans ce cadre, on peut considérer des attributs simples : niveau de gris (ex: imagerie médicale et scène d'intérieur [BSG94] ou multiples : texture [Bou93], couleur).

Dans le cadre du recalage 3D, il s'agit de considérer l'ensemble des appariements possibles comme espace de travail et non plus celui des attributs directement caractéristiques de l'image. Dans ce contexte, les classes significatives sont lourdement perturbées par un nuage de données parasites. C'est la raison qui nous amène à tenir compte d'avantage, dans notre démarche, de l'aspect modélisation du bruit.

De fait, nous nous sommes intéressés à une méthode de classification proposée par Davé [Dav91] où une classe supplémentaire nommée 'classe bruit' est introduite pour récolter tous les points parasites qui perturbent l'estimation des classes. La 'classe bruit' est définie par un prototype qui est à égale mesure de similarité de tous les points de l'espace des caractéristiques. L'algorithme est contrôlé avec un paramètre λ qui indique l'importance numérique des données parasites. Nous avons testé la méthode, qui donne d'excellents résultats quand le rapport SNR du nombre

de points à classer sur le nombre total est supérieur à 10%. Malheureusement, dans le cadre du recalage 3D, le nombre de points parasites est généralement beaucoup plus important (voir la formule de SNR dans la section 2.1). Sans l'ajout d'informations supplémentaires, l'algorithme est incapable de départager les données pertinentes du nuage de points aberrants, sauf dans les situations les plus simples.

2.2.2 Intégration de connaissances dans la classification floue

Comme la proportion de déplacements non valides est très importante, il faut augmenter l'efficacité de la classification. Pour cela, nous proposons l'introduction de connaissances géométriques sur les facettes traduites par un facteur de confiance C_k associé à chaque point de l'espace des transformations (équation (1)). L'idée est de pondérer la participation de chaque point de l'espace de classification par ce coefficient. Nous minimisons donc la fonctionnelle suivante :

$$J_m = \sum_{i=1}^c \sum_{k=1}^n C_k u_{ik}^m d^2(x_k, v_i) \quad (3)$$

où $d(x_k, v_i)$ est la distance entre le point x_k et le prototype v_i de la classe i . u_{ik} est le degré d'appartenance du point k à la classe i . u_{ik} est toujours positif et sa somme sur toutes les classes est égale à 1. Cette contrainte $\sum_{i=1}^c u_{ik} = 1$ force la méthode de classification à expliquer la totalité des données par c sous-ensembles flous (voir [Bez81] pour des explications plus complètes sur cet aspect des méthodes de classification floues).

L'algorithme de classification se déduit de la fonctionnelle (3) qui peut être résumé de la manière suivante :

- **étape 1** : Fixer le nombre de classes c et le facteur flou m . Calculer la position initiale des centres v_i des classes en utilisant l'algorithme des FCM, par exemple. Spécifier le paramètre λ relié à l'étendue du bruit.
- **étape 2** : Générer une partition en utilisant l'équation suivante du degré d'appartenance :

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}}$$

- **étape 3**: Calculer les nouveaux centres des classes en utilisant l'équation suivante (pour $i = 1$ à $c - 1$):

$$v_i = \frac{\sum_{k=1}^n C_k u_{ik}^m x_k}{\sum_{k=1}^n C_k u_{ik}^m}$$

- **étape 4**: S'arrêter quand la partition est stable, sinon retourner à l'étape 2.

x_k est un élément de l'ensemble de n points dans \mathbb{R}^p ($p = 6$ dans notre cas) et le facteur de confiance associé est C_k . m est le facteur flou généralement fixé à 1.5 dans nos expériences. d_{ik} est défini comme la distance euclidienne entre le point x_k de l'ensemble des données et le centre v_i de la classe, pour $i = 1$ à $c - 1$. Pour la classe de bruit, la définition de la distance est :

$$d_{ck}^2 = \frac{\lambda}{n(c-1)} \sum_{i=1}^{c-1} \sum_{l=1}^n d_{il}^2 \quad (4)$$

La convergence de la méthode de classification vers un minimum local, dans ce cadre étendu, est conservée (voir l'annexe 6 pour des compléments sur la démonstration).

L'introduction du coefficient de confiance permet de préserver la proportion des points valides parmi les points parasites et d'augmenter ainsi la qualité de la classification. Plus généralement, il est possible grâce à cette mesure de représenter des connaissances supplémentaires associées à chaque point de l'espace de classification, quelque soit la méthode utilisée.

2.3 Raffinement itératif du recalage

A ce stade, on dispose de $c - 1$ hypothèses intéressantes de déplacements qui recalent des parties des deux jeux de données 3D. La première étape utilise des facettes qu'elles supposent non occultées pour faire l'appariement. Les déplacements obtenus par la deuxième étape peuvent donc être légèrement biaisés par des appariements de faces occultées. Il est donc nécessaire d'avoir une étape de recalage plus fin au niveau des points pour traiter les cas d'occultation des faces. Dans ce contexte, une méthode itérative de recalage avec mise en correspondance explicite est bien adaptée. Lors de cette étape, l'algorithme travaille avec les points de contour

des faces, pour avoir même en présence de faces occultées une bonne précision du recalage.

La méthode itérative de recalage ICP [BM92] calcule à chaque pas une mise en correspondance de chaque point du premier ensemble de données avec le point le plus proche de l'autre ensemble à positionner. Ces correspondances permettent le calcul d'un déplacement optimal au sens des moindres carrés. L'algorithme itère les deux étapes appariement/recalage jusqu'à ce que l'erreur résiduelle soit stable. Cette procédure converge systématiquement vers le plus proche minimum local.

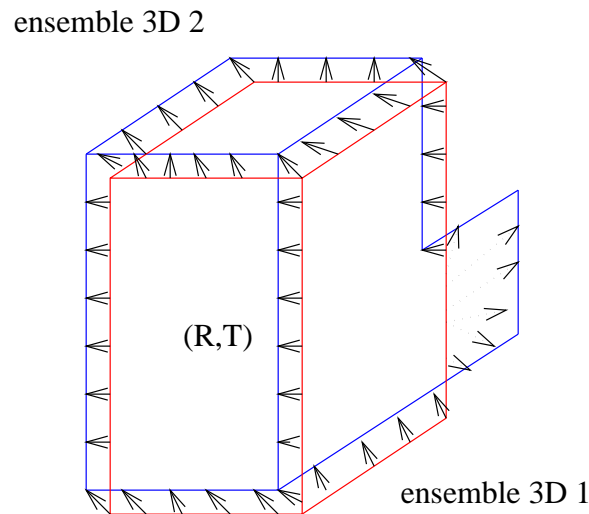


FIG. 5 - L'appariement est réalisé avec le plus proche voisin, et suivant la longueur du vecteur déplacement, la correspondance est prise ou n'est pas prise en compte.

La méthode ICP brute n'est pas adaptée à notre problème, car elle nécessite des données qui se recouvrent complètement. Pour remédier à cela [KVB93, Zha94], les mises en correspondance sont analysées statistiquement à chaque étape de l'algorithme pour ne retenir que les plus pertinentes au sens de la longueur. Les faux appariements sont alors en grande partie ignorés, ce qui permet le recalage avec des ensembles 3D se recouvrant partiellement (figure 5). Lors de cette analyse statistique, un paramètre P est introduit, qui indique un ordre de grandeur du pourcentage de points appariables (% du recouvrement).

L'étape de recalage fin est appliquée sur chaque hypothèse de déplacement fourni par la classification. Aussi, la comparaison des résidus après minimisation permet de trier les recalages les plus intéressants. Bien que les méthodes de type ICP soient généralement assez coûteuses en temps de calcul, le résultat souhaité s'obtient dans notre cas assez rapidement car la position initiale est suffisamment proche de la solution exacte.

2.4 Conclusion

La mise en correspondance, le recalage grossier puis fin, forment les trois étapes de la procédure de recalage robuste. Souvent, les reconstructions sont très partielles à cause du point de vue, ou des défauts du processus d'analyse d'image, par exemple. Pourtant, la méthode proposée à l'avantage de permettre le recalage 3D d'un modèle d'objet isolé sur ce type de reconstructions partielles. Ainsi, nous avons pu obtenir sans difficultés des résultats sur des ensembles qui possèdent une partie commune inférieure à 20% de la taille du plus grand ensemble à recalier.

3 Résultats

La méthode présentée a été intégrée à un système d'analyse de scènes d'intérieur à partir d'images stéréoscopiques (voir figure 6). Notre processus d'analyse consiste en quatre étapes [JMN⁺96]:

- Une calibration géométrique [TG95] et radiométrique [Tar96] des caméras.
- Une segmentation de chaque image en régions [RG91, AMG93, BSG94].
- Une mise en correspondance 2D des régions segmentées entre les deux images [SVCG89, Ran92].
- Une reconstruction des facettes 3D à partir des paires de régions que l'on suppose être les images d'une même facette [TV95].

Notre méthode de recalage a été validée complètement sur les objets synthétiques où le déplacement à obtenir est parfaitement connu. Des tests sur les données réelles ont permis de valider la méthode sur ce type de données, dans la mesure où le déplacement exact est inconnu.

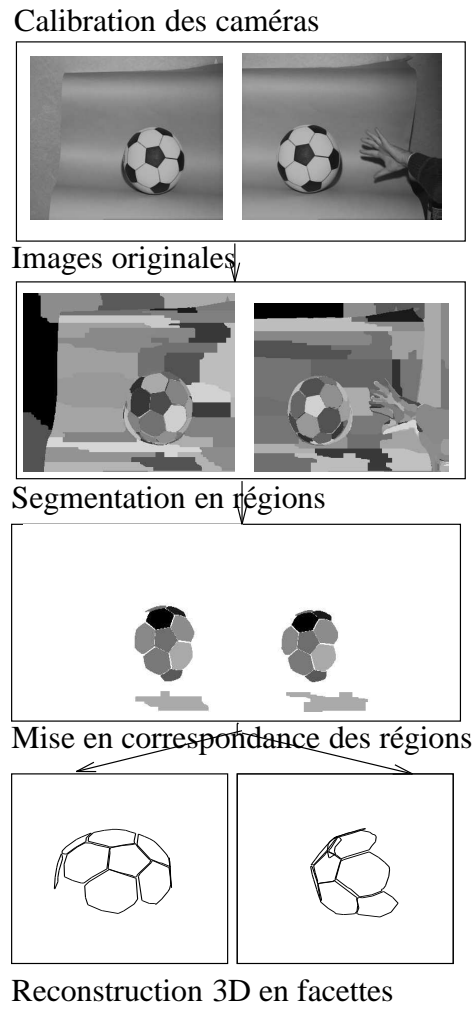


FIG. 6 - *Le processus d'analyse stéréoscopique.*

3.1 Réglage des paramètres

Notre méthode suppose la connaissance d'un certain nombre de paramètres qui sont les suivants :

- I** : écart non significatif pour comparer la longueur des axes d'inertie. I indique la tolérance observable sur les données d'entrée du recalage. Ainsi, l'examen de la précision obtenue par les algorithmes ou capteurs qui produisent ces données 3D permet de régler I une fois pour toutes.
- λ : un paramètre lié au rapport du nombre d'appariements justes de facettes sur leur nombre total. C'est le seul paramètre réellement critique du point de vue de la qualité des résultats. Expérimentalement, nous avons constaté que la valeur de SNR permet d'avoir un bon ordre de grandeur de λ (à un facteur constant près).
- c : le nombre de classes, ou de déplacements potentiels à estimer. Le nombre de classes c doit être choisi après l'analyse des symétries de l'ensemble 3D à recalculer qui peuvent conduire à l'existence de plusieurs solutions également équivalentes (voir 3.2). Il faut aussi tenir compte d'éventuelles apparitions multiples de l'objet à recalculer dans l'ensemble à interpréter. Néanmoins, ce paramètre peut être largement sous évalué sans difficulté : seul un sous-ensemble des solutions possibles est alors obtenu.
- P** : la proportion de points non appariables lors du recalage local fin. Ce paramètre est lié à λ . Ce pourcentage P peut être largement sous évalué. Il correspond en fait à une borne supérieure du pourcentage de mises en correspondance de points qui peuvent s'avérer fausses, dans l'étape de recalage fin.

En conséquence, la méthode de recalage est simple d'utilisation puisque contrôlée par le paramètre λ et le nombre $c - 1$ de solutions que l'on désire obtenir. Ceci permet d'envisager une automatisation complète du processus de recalage.

3.2 Les objets symétriques

Les situations où l'ensemble à recalculer présente des symétries ou apparaît plusieurs fois sont généralement difficiles à gérer avec les méthodes classiques. Notre

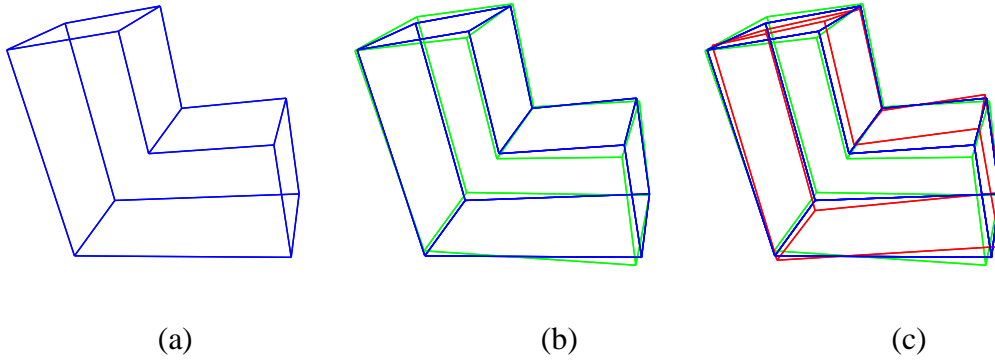


FIG. 7 - (a) Objet 3D de synthèse. (b) Une solution de la phase du recalage approché qui possède deux solutions également possibles du fait des symétries de l'objet (c).

méthode présente l'avantage de manipuler l'ensemble des solutions possibles de manière unifiée. En effet, il permet de choisir le nombre de classes c en fonction du nombre de recalages que l'on désire obtenir. La classification floue permet d'éviter la redondance des solutions en contraignant les classes à recouvrir au mieux l'ensemble des déplacements possibles.

	rot_x	rot_y	rot_z	tr_x	tr_y	tr_z
classe1	0.005	0.882	0.017	0.017	-0.002	0.007
juste 1	0.000	1.000	0.000	0.000	0.000	0.000
classe2	2.919	-0.012	0.854	0.006	0.016	-0.016
juste 2	3.014	0.000	0.884	0.000	0.000	0.000

TAB. 1 - Déplacements obtenus par la phase de recalage grossier et les valeurs effectives de la figure 7. La rotation est représentée par les composantes de son axe normalisé multiplié par son angle en radian.

Par exemple, dans la figure 7, l'objet synthétique à recaler est identique à l'objet de référence, à un déplacement rigide près. Du fait des symétries de cet objet, deux

recalages sont possibles qui sont effectivement bien retrouvés. Le tableau 1, où sont comparés les déplacements réels et estimés par la phase de recalage global ($c = 3$, $\lambda = 0.035$), en témoigne. Le résultat du recalage fin aboutit à une superposition parfaite avec l'objet de la figure 7(a).

Plus généralement, la distance moyenne entre les parties communes des deux modèles après le recalage fin, fourni par cette dernière étape, donne une valeur pour trier les solutions suivant leur qualité.

3.3 Utilisation de la couleur

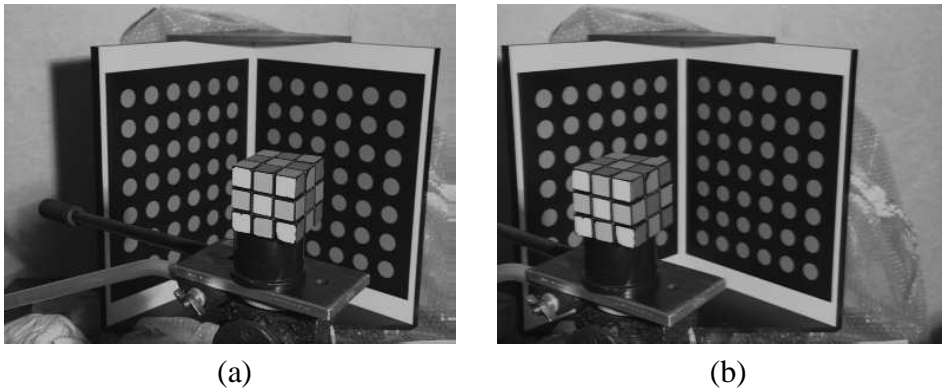


FIG. 8 - (a)(b) paire stéréoscopique traitée pour obtenir la reconstruction de figure 9(a).

Une facette ne possède pas que des caractéristiques propres de nature géométrique. Elle peut aussi posséder une couleur moyenne. Cette information de couleur, décrite par les composantes de couleurs (r, v, b) normalisées par exemple, peut être intégrée au facteur de confiance (équation 1). Nous l'avons fait de la manière suivante :

$$C_{ij} = \min\left(1, \frac{I}{|\alpha_i - \alpha_j|}\right) \min\left(1, \frac{I}{|\beta_i - \beta_j|}\right) \\ \min\left(1, \frac{G}{|r_i - r_j|}\right) \min\left(1, \frac{G}{|g_i - g_j|}\right)$$

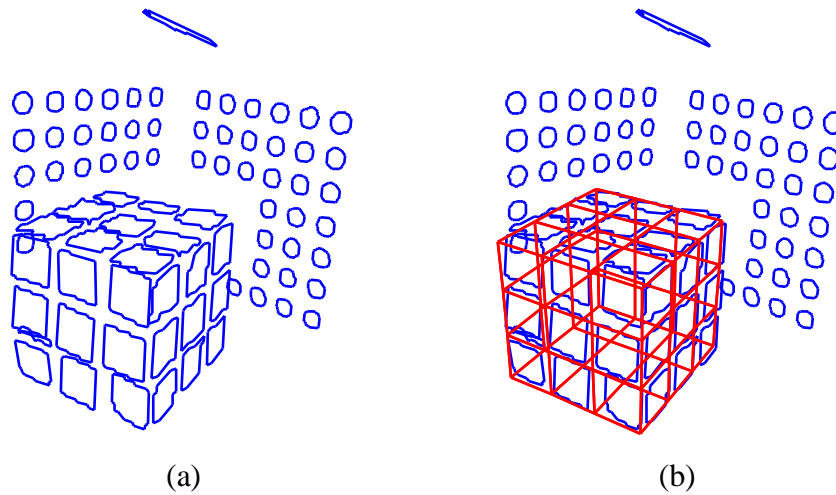


FIG. 9 - (a) Vue de la reconstruction 3D obtenue par analyse sur la paire de la figure 8. (b) Le modèle parfait du rubik cube est recalé sur la reconstruction.

où G est l'erreur moyenne pour la comparaison des couleurs. G est déterminé à partir du pouvoir séparateur des couleurs du système d'analyse.

3.4 Recalage avec plusieurs objets

La méthode présentée, vu sa robustesse, permet de faire du recalage avec plusieurs objets. A fin d'illustration, le résultat du processus d'extraction de facettes 3D sur les images originales (a) et (b) de la figure 10 est visible dans la figure 11(a)(b). Cette scène contient deux objets principaux : une mire et un ballon, dont nous avons le modèle 3D complet. Ces deux modèles ont donc été recalés avec succès par notre méthode de recalage floue.

Dans ce cas là, les résultats obtenus en 3D sont suffisamment précis pour que le modèle recalé se reprojette bien sur la paire stéréoscopique. Il n'est donc pas nécessaire d'affiner le recalage par une méthode de 3D/2D qui minimise l'erreur directement entre le modèle et les images.

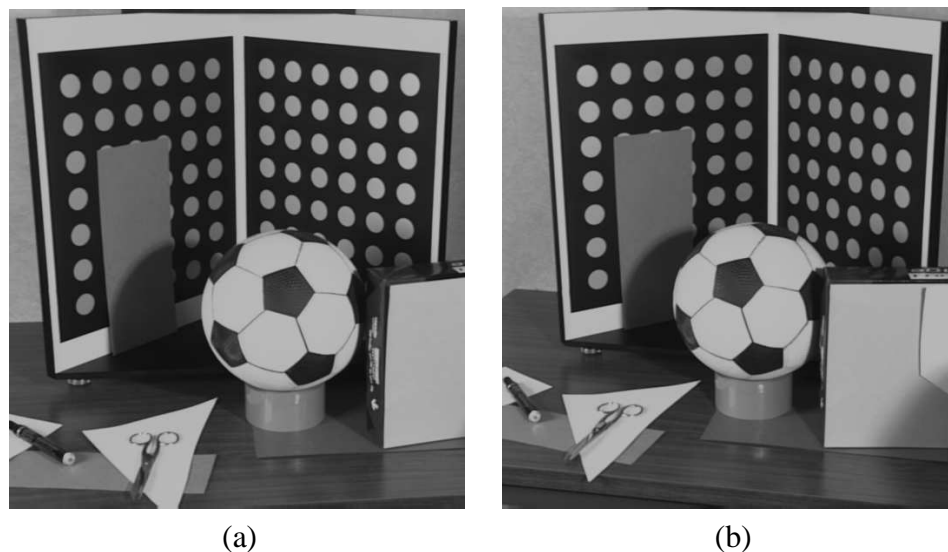


FIG. 10 - (a)(b) Images originales de la scène. Les vues des résultats de reconstruction 3D obtenus avec nos algorithmes de stéréovision sont exposées dans la figure 11(a)(b).

3.5 Validation par rétroprojection

Une fois les principaux objets de la scène recalés, il est possible de reprojeter une des images originales de la paire stéréoscopique sur les modèles 3D recalés pour vérifier l'exactitude du recalage. Nous avons utilisé l'image comme une diapositive dans le programme de rendu d'images par lancer de rayons Rayshade [FvDFH90] sur les différentes scènes ici utilisées. Quelques images de synthèse générées sont visibles sur la figure 12.

4 La similarité entre données 3D

Suite aux résultats obtenus, il se révèle indispensable d'avoir un outil de comparaison des données 3D. Cet outil est en effet très utile pour valider notre méthode de recalage. De plus, il permet l'évaluation objective de nos résultats par rapport aux autres approches.

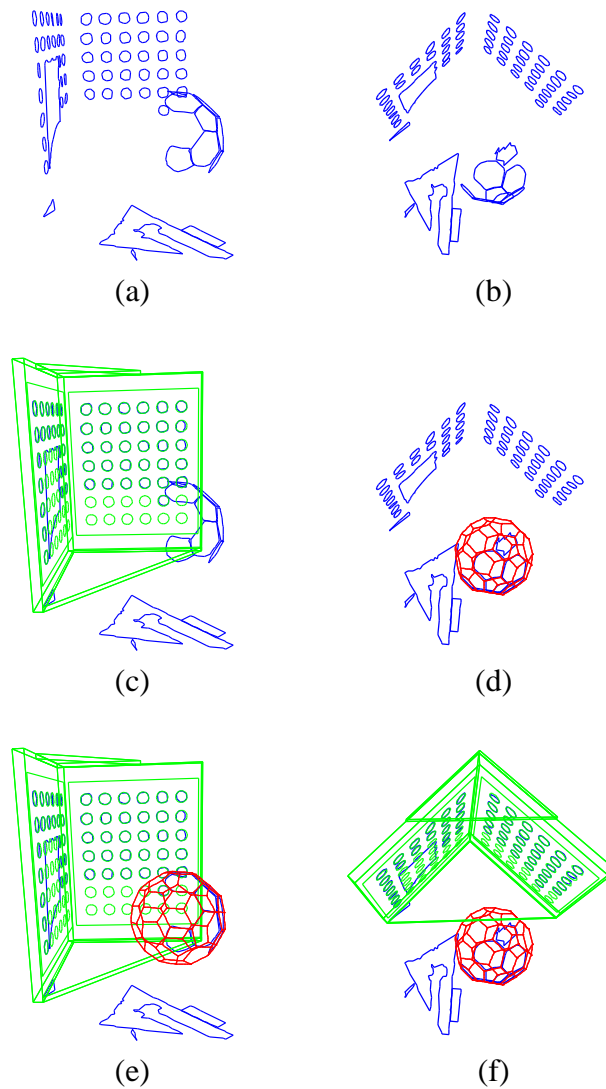


FIG. 11 - (a) et (b) sont deux points de vue de la reconstruction 3D obtenue à partir de la paire de la figure 10. Dans (c), la mire est recalée sur la reconstruction. Dans (d), seul le ballon est recalé, alors que (e) et (f) montrent les deux objets recalés. Ces résultats ont été obtenus par l'application complète des trois étapes de la méthode pour avoir un recalage très précis.

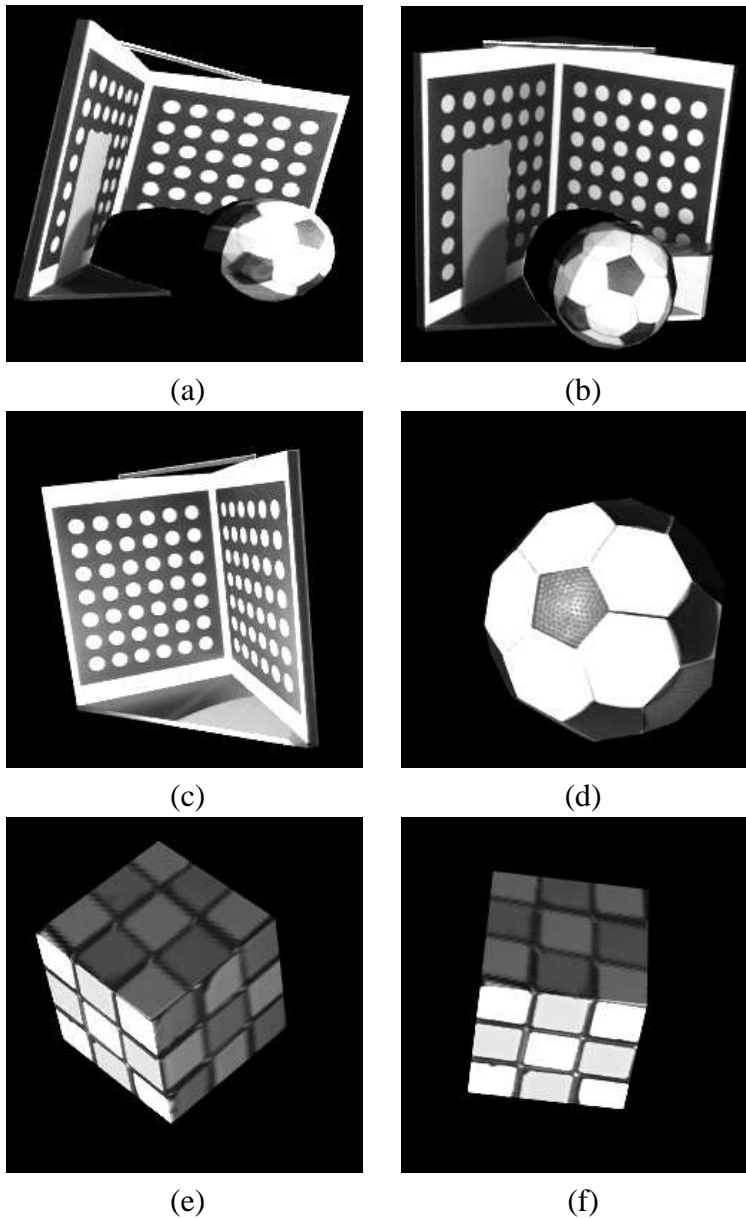


FIG. 12 - (a)(b) Point de vue de synthèse des modèles recalés (voir figure 11(e)(f)) sur lesquels l'image originale 10(a) est projetée. (c) Point de vue synthétique de la mire de calibration seule. (d) Point de vue synthétique du ballon obtenu après l'analyse de la paire de la figure 6. (e)(f) Point de vue de synthèse sur la scène du rubik cube de la figure 9(b) sur lequel est projetée l'image de la figure 8(a).

4.1 Présentation de l'outil de mesure

La définition de cet outil pose le problème suivant : en présence de deux ensembles de données 3D, par quels critères estimer la ressemblance de deux jeux de données? Comme les deux ensembles de données 3D que nous utilisons ne peuvent se recouvrir que partiellement, il est nécessaire de sélectionner les parties des deux ensembles qui peuvent être mises en correspondance. En conséquence, chaque ensemble 3D est divisé en deux parties dont l'une ne peut pas être mise en correspondance. Sur la portion des données qui ne recouvre pas le modèle, la seule mesure intéressante est sa taille relative. Par contre, sur la partie commune, les deux mesures importantes sont le biais et l'écart type des déplacements entre les données mises en correspondance.

On suppose que les ensembles 3D sont constitués de points, puisqu'il est toujours possible d'échantillonner avec des points des géométries linéiques et surfaciques. Nous avons choisi, pour pouvoir estimer les indices de similarité, de faire la mise en correspondance de chaque point 3D du premier ensemble dit de référence avec le point le plus proche en distance de l'autre ensemble comme dans la section 2.3 (voir figure 5).

Ensuite, une classification est réalisée sur les déplacements des points en correspondance entre les deux ensembles. Un algorithme de segmentation floue permet de séparer de manière adaptative, sans difficultés (comme en témoigne par exemple l'ampleur du pic de la figure 13), la part des données en recouvrement de celle qui est en fausse correspondance. L'algorithme de classification avec bruit génère deux classes : la première récolte les déplacements faibles et la classe de bruit le reste des déplacements 3D.

Une fois l'ensemble 3D segmenté comme dans la figure 14, il est aisé de calculer sur la première classe des critères statistiques sur les déplacements. Ainsi, le centre de cette classe s'interprète comme le biais moyen du déplacement entre la partie commune des deux ensembles 3D. Le rayon de la classe est la distance moyenne entre les parties en correspondances. Enfin, la taille relative de la partie sans mise en correspondance est celle de la classe rejet.

4.2 Résultats

D'après le tableau 2, nous avons obtenu entre la reconstruction 3D de la figure 11(a)(b) et la mire recalée, une erreur totale de 2.9 mm. Cette scène présente

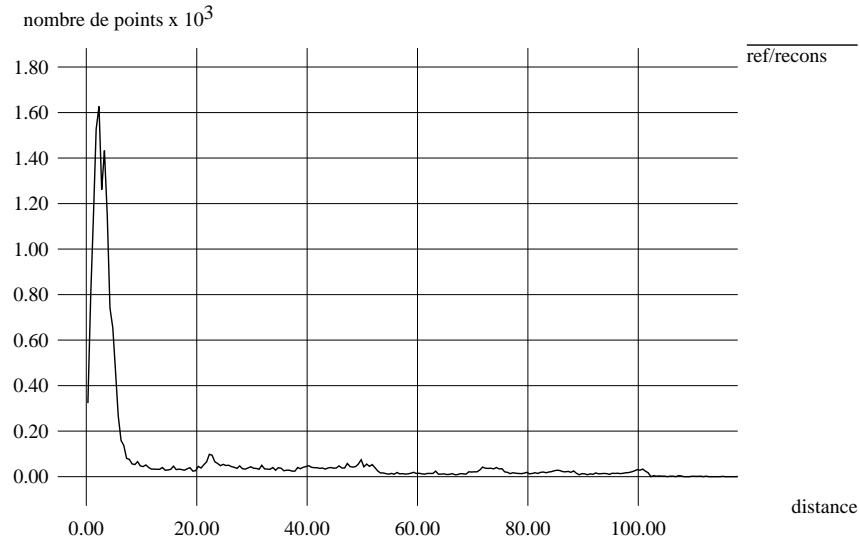


FIG. 13 - *Histogramme des longueurs en mm des déplacements entre les points appariés entre les deux ensembles 3D. Dans cette exemple, l'objet de référence est celui de la figure 14 et le deuxième ensemble et la reconstruction 3D visible dans la figure 11(a)(b).*

taille de la partie sans correspondance	32.1%
biais moyen sur la partie en correspondance	1.78 mm
écart type sur la partie en correspondance	1.11 mm

TAB. 2 - *Biais et erreur sur la partie en recouvrement ainsi que la taille de la partie sans correspondance des points de la mire de la figure 11(d)(e)(f).*

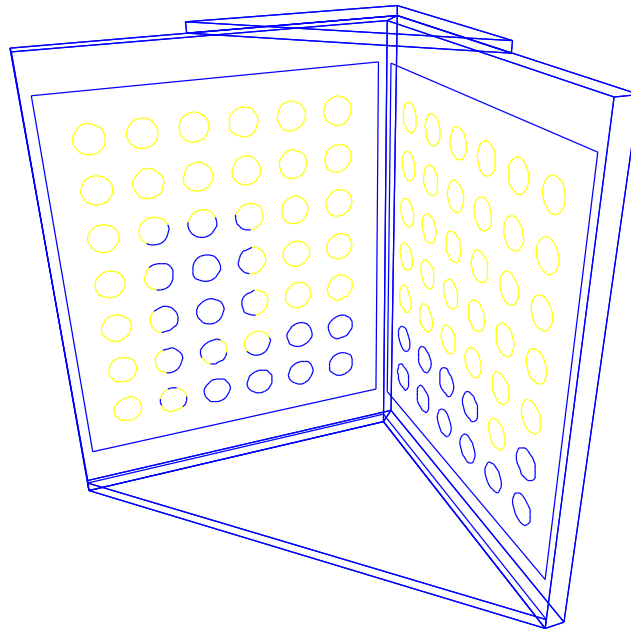


FIG. 14 - *Segmentation de l'ensemble de référence en sa partie commune avec la reconstruction (en gris) et sa partie non appariable. La reconstruction 3D associée est visible sur la figure 11(a)(b).*

une profondeur de 1200 mm avec une base stéréoscopique de 340 mm. Théoriquement, l'erreur de reconstruction est donc de 1.2 mm de face et 9 mm selon l'axe de la caméra [Aya88]. On peut grossièrement attribuer l'écart type aux erreurs de reconstruction 3D et donc associer le biais à l'erreur de recalage. En conséquence, le paramètre I du critère (1) est fixé autour de 2 mm.

5 Conclusion

Nous proposons une méthode de recalage de données 3D constituée de trois étapes qui combine aussi bien les avantages complémentaires des méthodes à appariement explicite (grands déplacements), et à appariement implicite (précision et

occlusions au niveau du point). Les deux premières étapes constituent une méthode par appariement explicite robuste qui est fondée sur un algorithme de classification floue. Les résultats obtenus mettent en évidence l'intérêt des approches par classification floue pour la résolution des problèmes en vision par ordinateur. De plus, il nous semble intéressant d'examiner l'apport des approches possibilistes dans ce contexte.

Notre méthode permet d'obtenir un recalage suffisamment précis pour faire de l'interprétation de données 3D à partir de modèles et de la concaténation de reconstructions 3D à partir de points de vue différents. Les ensembles sont recalés de manière rigide, mais une simple extension de notre approche permettra de combiner une transformation rigide avec un changement d'échelle.

Enfin, le résidu obtenu après recalage fin, sur les différents modèles géométriques d'objets qui peuvent se trouver dans la scène, permet d'envisager l'utilisation de notre méthode non seulement pour la reconnaissance 3D d'un modèle dans la reconstruction, mais aussi pour la reconnaissance à partir de bibliothèques d'objets.

6 Annexe

Nous présentons dans cette section la ligne directrice de la démonstration de la convergence vers un minimum local de l'algorithme FCM, lorsque que la valeur de confiance C_k est introduite comme cela est proposé dans la section 2.2.2. Cette démonstration est une adaptation de la convergence de FCM [Bez81, BHST87] avec la même condition sur le facteur flou : $m > 1$. Une extension similaire peut être faite à partir de la preuve de l'algorithme de Davé.

Soit la matrice $U = (u_{ik})$ et le vecteur $V = (v_i)$. Nous voulons minimiser la fonctionnelle suivante :

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n C_k u_{ik}^m d^2(x_k, v_i) \quad \text{avec} \quad \sum_{i=1}^c u_{ik} = 1 \quad \text{et} \quad u_{ik} \geq 0 \quad (5)$$

où d est la distance Euclidienne. La démonstration de la convergence théorique est fondée sur l'application du théorème de Zangwill qui assure la convergence après un nombre infini d'itérations comme dans une méthode de descente de gradient ou de Newton.

Premièrement, nous fixons V . Puisque les colonnes de la matrice U sont indépendantes, nous déduisons que :

$$\min(J_m(U)) = \sum_{k=1}^n C_k \min_{\sum_{i=1}^c u_{ik}=1} \left(\sum_{i=1}^c u_{ik}^m d^2(x_k, v_i) \right)$$

Alors, pour chaque terme, il faut faire une minimisation sous contrainte. En appliquant les multiplicateurs de Lagrange, nous obtenons le minimum sous la forme suivante, après calcul :

$$\frac{u_{ik} = 1}{\sum_{j=1}^c \left(\frac{d_{ijk}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (6)$$

Deuxièmement, nous fixons U . Soit $\langle \cdot \cdot \rangle$ le produit scalaire. Alors, il faut minimiser le terme suivant :

$$J_m(V) = \sum_{i=1}^c \sum_{k=1}^n C_k u_{ik}^m \langle x_k - v_i \cdot x_k - v_i \rangle$$

Pour chaque i , il est nécessaire que les dérivées $J'_m(v_i, g)$ suivant chaque vecteur unité $g \in \mathbb{R}^c$ soient nulles :

$$\begin{aligned} J'_m(v_i, g) &= -2 \sum_{k=1}^n C_k u_{ik}^m \langle x_k - v_i \cdot g \rangle = 0 \quad \forall g \\ &\Leftrightarrow \sum_{k=1}^n C_k u_{ik}^m (x_k - v_i) = 0 \end{aligned}$$

D'où, nous déduisons que :

$$v_i = \frac{\sum_{k=1}^n C_k u_{ik}^m x_k}{\sum_{k=1}^n C_k u_{ik}^m} \quad (7)$$

Le prototype v_i peut être interprété comme le centre de gravité de l'ensemble $(x_k)_k$, où chaque point est pondéré par $C_k u_{ik}^m$. Il faut noter que les équations (6) et (7) permettent d'écrire les étapes 2 et 3 qui sont au coeur de l'algorithme de classification décrit dans la section 2.2.2.

Une interprétation duale de la fonctionnelle (5) est la suivante :

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n w_{ik}^m d^2(x_k, v_i)$$

où $w_{ik} = \sqrt[m]{C_k} u_{ik}$. Cette formule est la même dans l'algorithme FCM, où les contraintes sur w_{ik} ne sont pas $\sum_{i=1}^c w_{ik} = 1$ mais $\sum_{i=1}^c w_{ik} = \sqrt[m]{C_k}$. Ainsi, nous pouvons interpréter notre généralisation de l'algorithme de classification floue introduisant la pondération par la confiance C_k sur chaque point, comme une méthode de classification avec une contrainte non homogène sur les degrés d'appartenance. Ceci révèle un lien intéressant avec les approches possibilistes de la classification [KK93].

Références

- [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.
- [Aya93] Ayache (N.). – Computer vision applied to 3D medical imagery: Results, trends and future challenges. *In: Proceedings of the 6th Symposium on Robotics Research*. – MIT Press, also INRIA Tech. Report #2050, 1993.
- [Bez81] Bezdek (J.C.). – *Pattern Recognition with Fuzzy Objective Function Algorithms*. – Plenum Press, 1981.
- [BHST87] Bezdek (J.C.), Hathaway (R. J.), Sabin (M. J.) et Tucker (W. T.). – Convergence theory for fuzzy c-means: Counterexamples and repairs. *IEEE Trans. Systems, Man, and Cybernetics*, vol. 17, n° 5, 1987, pp. 873–877.

- [BM92] Besl (Paul J.) et McKay (Neil D.). – A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, n° 2, February 1992, pp. 239–256.
- [Bou93] Boujemaa (N.). – *Modélisation floue de l'incertitude pour la segmentation d'images*. – Thèse de PhD, Université Paris-V, 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*.
- [Dav91] Dave (R.N.). – Characterization and detection of noise in clustering. *Pattern Recognition Letters*, vol. 12, 1991, pp. 657–664.
- [FH83] Faugeras (O.) et Hebert (M.). – The representation, recognition and positioning of 3d shapes from range data. In : *Proceedings of the 8th International Conference On Artificial Intelligence*, pp. 996–1002. – Karlsruhe, BRD, August 1983.
- [FvDFH90] Foley (J. D.), van Dam (A.), Feiner (S. K.) et Hughes (J. F.). – *Computer Graphics: Principles and Practices (2nd Edition)*. – Addison Wesley, 1990.
- [GA92] Guézic (A.) et Ayache (N.). – Smoothing and matching of 3D-space curves. In : *Proceedings of the Second European Conference on Computer Vision*. – Santa Margherita Ligure, Italy, May 1992.
- [HH94] Hoogs (A.) et Hackett (D.). – Model-supported exploitation as a framework for image understanding. In : *Proceedings of the ARPA IU Workshop*.
- [JMN⁺96] Jancène (P.), Meilhac (C.), Neyret (F.), Provot (X.), Tarel (J.-P.), Vézien (J.-M.) et Vérroust (A.). – Réalité enrichie par synthèse. In : *10ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*. – Rennes, France, 1996.
- [KK93] Krishnapuram (R.) et Keller (M.). – A possibilistic approach to clustering. *IEEE Trans. Fuzzy Systems*, vol. 1, n° 2, May 1993, pp. 98–110.

- [KPLR91] Kamgar-Parsi (B.), Lones (J.) et Rosenfeld (A.). – Registration of multiple overlapping range images: Scene without distinctive features. *Pattern Analysis And Machine Intelligence*, vol. 13, n° 9, September 1991, pp. 857–870.
- [KVB93] Koivunen (V.), Vézien (J.-M.) et Bajcsy (R.). – Procedural cad-models from range data. In : *Proceedings of Towards World Class Manufacturing '93*.
- [Ran92] Randriamasy (S.). – *Segmentation descendante coopérative en régions de paires d'images stéréoscopiques*. – Thèse de PhD, Université Paris-IX Dauphine, 1992.
- [RG91] Randriamasy (S.) et Gagalowicz (A.). – Region based stereo matching oriented image processing. In : *Proceedings of Computer Vision and Pattern Recognition*. – Maui, Hawaii, June 1991.
- [RH91] Rigoutsos (I.) et Hummel (R.). – Robust similarity invariant matching in the presence of noise: a data parallel approach. In : *Proceedings of the 8th Israeli Conference on Artificial Intelligence and Computer Vision*.
- [Sto87] Stockman (G.). – Object recognition and localization via pose clustering. *Computer Vision, Graphics and Image Processing*, vol. 40, n° 2, 1987, pp. 361–387.
- [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>.
- [Zha94] Zhang (Z.). – Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, vol. 13, n° 2, 1994, pp. 119–152. – also INRIA Tech. Report #1658.



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